El Principio de Relatividad

y el problema del conocimiento

Segunda edición con correcciones y ampliaciones

Roberto Suárez Ántola

ISBN: 978-9974-94-596-8

Suárez Ántola, Roberto

El principio de la relatividad y el problema del conocimiento.

Segunda edición corregida y ampliada.

2a ed. - Montevideo, 2019.

ISBN 978-9974-94-596-8

Materia: 530.15 - Física matemática PHR - Física de la relatividad

Datos de la primera edición del libro:

Suárez Ántola, Roberto

El principio de la relatividad y el problema del conocimiento.

1a ed. - Buenos Aires: Dunken.

ISBN 978-987-02-5680-9

1. Física. 2. Epistemología. 3. Matemática. I. Título

CDD 530

Contenido

Prólogo a la segunda edición 9

Prólogo a la primera edición 9

1. Introducción 12

1.1 Mediciones de longitudes y tiempos 13

1.2 Geometrías 14

1.3 Ciencia y principio de causalidad 15

1.4 Principios de relatividad y teorías de la relatividad 17

1.5 Ciencia, filosofía y matemática 18

2. La mecánica newtoniana de partículas y su principio de relatividad 26

2.1 Marcos de referencia inerciales: reglas y relojes newtonianos 28

2.2 Las leyes de la mecánica 29

2.2.1 Primera, segunda y tercera ley de Newton 30

2.2.2 Conservación de la cantidad de movimiento y del momento angular de un sistema de partículas. Centro de masas. 31

2.3 El principio de relatividad de la mecánica clásica 32

2.4 Atracción gravitatoria 37

2.6 Mecánica de medios continuos y óptica ondulatoria 38

2.7 Mecánica clásica, causalidad eficiente, determinismo y libre albedrío 39

3. Campos y principio de contigüidad 44

3.1 Campos y energía en mecánica newtoniana de partículas 45

3.2 Campos en mecánica de medios continuos y electrodinámica: la versión fuerte del principio de contigüidad. 46

3.3 Campos en la teoría de los procesos de transporte de calor y masa: la versión débil del principio de contigüidad. 47

3.4 El postulado de posibilidad de descripción continua de los fenómenos físicos 48

3.5 Reconsideración del principio de causalidad 51

4. La electrodinámica clásica 53

4.1 Electro-estática, magneto-estática y electrodinámica pre-relativista 54

4.2 La electrodinámica de los cuerpos en movimiento 57

4.3 La teoría clásica del electrón 59

5. Las ciencias físicas al final del siglo XIX 62

5.1 La mecánica newtoniana como súper-teoría física..... 62

5.2 El problema del éter 63

5.3 La ecuación de ondas clásica: la tridimensionalidad del espacio físico y la evolución de seres inteligentes 68

6. Los conceptos de tiempo y espacio en la teoría restringida de la relatividad..... 71

6.1 El principio de relatividad de Einstein, las transformaciones de Lorentz y sus consecuencias cinemáticas 72

6.2 Homogeneidad del tiempo, homogeneidad e isotropía del espacio y principio de relatividad de Einstein..... 78

6.3 Formulación de la teoría de la relatividad independiente del postulado de constancia de la velocidad de la luz: la constante de estructura del tiempo-espacio 81

7. La interrelación entre el tiempo y el espacio: el continuo de tetra-dimensional de Minkowski y la causalidad en la teoría restringida de la relatividad..... 85

7.1 El grupo de transformaciones de Lorentz y la crono-geometría de Minkowski 87

7.2 La relatividad del espacio físico y sus consecuencias. Líneas y tubos de universo. 88

7.3 El tiempo en la teoría restringida de la relatividad..... 91

7.4 Causalidad y determinismo en el marco de la teoría de relatividad 93

8. Física relativista 97

8.1 Dinámica relativista y leyes de conservación..... 98

8.1.1 Dinámica de una partícula..... 98

8.1.2 Sistemas de partículas y leyes de conservación 101

8.1.3 Transformación de la fuerza al cambiar de marco inercial. Acción y reacción 103

8.2 Electrodinámica, relatividad y cuantos 104

8.3 Partículas y corpúsculos: fotones, campos gravitatorios y relojes 105

8.4 Invariancia de la fase de una onda..... 107

8.4.1 Efecto Doppler y aberración relativista de la luz 107

8.4.2 La apariencia visual de objetos que se mueven con velocidades cercanas a la de la luz en el vacío 109

8.5 Cuadrivectores, dinámica relativista y campo electromagnético 111

8.6 La fórmula para el tiempo propio de los relojes acelerados 115

8.7 Algunos resultados relacionados con la mecánica de fluidos y la termodinámica relativistas 117

8 .8 Las relaciones epistemológicas entre las teorías de la física clásica luego de ser aceptada la teoría restringida de la relatividad 123

9 Espacios curvos 127

9.1 Teoría de superficies y geometrías intrínsecas 127

9.1.1 Curvas, superficies y curvaturas en E_3. 127

9.1.2 Resumen de algunos resultados sobre geometrías intrínsecas de superficies en E_3. 134

9.2 ¿Curvatura o fuerza? 138

9.3 Movimiento libre de una partícula newtoniana en una superficie 139

9.4 Espacios abstractos y variedades de Riemann 141

9.4.1 Resumen del aparato analítico de la geometría diferencial clásica. 141

9.4.2 Espacios euclidianos de n dimensiones 148

9.4.3 Variedades riemannianas de n dimensiones 150

9.4.4 Variedades lineales tangentes 151

9.4.5 Inmersión de variedades riemannianas en espacios euclidianos 153

9.4.6 Variedades 3-D con simetría esférica. Representación conforme en E_3. 157

9.5 Tensores y curvatura. 158

9.5.1 Tensores 158

9.5.2 Coeficientes de conexión, desplazamiento paralelo y derivadas covariantes 161

9.5.3 Curvas geodésicas 164

9.5.4 Curvatura de una variedad de n dimensiones 167

10. La gravitación y la teoría generalizada de la relatividad 174

10.1 El principio de equivalencia 174

10.2 El tiempo-espacio-materia y la gravitación 178

10.2.1 Covariancia 180

10.2.2 Consistencia y tanteos preliminares 181

10.2.3 Las ecuaciones del campo gravitatorio 183

10.2.4 La ley del movimiento en un campo gravitatorio 187

10.3 La relatividad restringida y la gravitación newtoniana como casos límite. Deducción de la conexión entre κ y G. 188

10.4 Las relaciones epistemológicas entre las teorías de la física clásica luego de ser aceptada la teoría generalizada de la relatividad 191

10.5 El significado de la teoría matemática de la relatividad 193

11. Solución de Schwarzschild y confirmaciones de la teoría de la gravitación 196

11.1 El cálculo del tensor métrico mediante el modelo de Lenz. Métrica de Schwarzschild. Singularidades de coordenadas y singularidades gravitacionales. 197

11.2 Gravitación y curvatura del continuo tiempo-espacio 200

11.3 Corrimiento al rojo en un campo gravitatorio 203

11.4 Movimiento de corpúsculos en un tiempo-espacio con la métrica de Schwarzschild: avance del perihelio de una órbita y deflexión de la dirección de propagación de la luz. 205

11.4.1 Elementos de cálculo de variaciones 205

11.4.2 Geodésica temporaloide 208

11.4.3 Geodésica nula 211

11.5 Confirmaciones de la teoría de la gravitación de Einstein 212

11.6 Óptica gravitacional: lentes, espejismos y materia oscura 214

12. El problema cosmológico 219

12.1 Aspectos históricos 220

12.2 El corrimiento del espectro de las galaxias y el horizonte de eventos 222

12.3 Modelos de Universo 226

12.3.1 El principio cosmológico. 228

12.3.2 Las ecuaciones de la relatividad generalizada y la constante cosmológica 229

12.3.3 El flujo de Hubble y el postulado de Weyl. 230

12.4 Modelos basados en la métrica de Friedmann-Lemaître-Robertson-Walker 232

12.4.1 Estudio estático de las geometrías a partir del término espacial del intervalo. Distancias propias y parámetro de Hubble. 235

12.4.2 Estudio a partir de la métrica completa. El modelo estándar Λ-CDM 238

12.4.3 Corrimiento al rojo cosmológico 246

12.5 El modelo cosmológico Λ-CDM y la teoría del Big Bang. 248

12.6 Modelo de Universo en mosaico. 254

13. La metagalaxia 255

13.1 Modelos de Metagalaxia 257

13.2 Estrellas 259

13.2.1 Enanas blancas 260

12.2.2 Estrellas de neutrones 264

13.3 Agujeros negros 267

14. Aspectos vinculados con la lógica de la investigación científica 277

14.1 La estructura de las teorías físico-matemáticas 278

14.2 El principio de relatividad y la teoría de la relatividad 283

14.3 Relatividad restringida 285

14.3.1 El problema de la velocidad límite 285

14.3.2 Simetrías geométricas y simetrías dinámicas 287

14.4 Relatividad generalizada: la teoría de la gravitación y algunos de sus problemas 288

14.4.1 Marcos de referencia locales y coincidencias puntuales 288

14.4.2 Algunas limitaciones de la teoría generalizada de la relatividad 290

14.5 Cosmología relativista 291

14.5.1 Dos métodos para responder al problema de la unicidad del Universo. 292

14.5.2 Tiempo cósmico, espacio cósmico, marcos inerciales y modelos cosmológicos... 294

15. Perspectivas psicológicas y sociales 298

15.1 Einstein y el mundo de Parménides. 299

15.2 El juego de ajedrez de la ciencia 302

15.3 El reconocimiento individual a la labor científica. 305

15.4 Especialización, genialidad y desmesura. 307

15.5 La sociedad post-moderna y la relatividad. 310

16. Epílogo: El principio de relatividad y el problema del conocimiento 317

16.1 Instrumentalismo, realismo y relatividad del movimiento. 318

16.2 Relatividad de las categorías del conocimiento 320

16.3 Ciencia, filosofía y concepciones del mundo. 322

17. Bibliografía 325

17.1 Temas de matemática 325

17.2 Temas de física. 325

17.3 Temas generales relacionados con relatividad y cosmología. 325

17.4 Temas de relatividad restringida y generalizada. 326

17.5 Temas de astronomía, astrofísica, relatividad general y cosmología física. 326

17.6 Temas de filosofía, historia y metodología de la ciencia. 326

Prólogo a la segunda edición

Esta es la segunda edición, corregida y ampliada, de un libro destinado a introducir al lector en los fundamentos fisicomatemáticos y epistemológicos de la teoría de la relatividad, tanto restringida como generalizada. Se consideran aplicaciones astrofísicas y cosmológicas de la teoría. Además, en los últimos capítulos se tratan algunos aspectos psicológicos y sociológicos de la investigación científica y se discute el principio de relatividad y algunas de sus consecuencias desde una perspectiva filosófica.

El texto de la primera edición fue terminado el 31 de julio de 2011 y editado a comienzos de 2012, en la ciudad de Buenos Aires. El texto de esta segunda edición tiene en cuenta las observaciones con el sistema de radiotelescopios EHT (en la sección 13.3) y los resultados de los nuevos experimentos sobre haces de neutrinos (en la sub-sección 14.3.1). En las secciones 8.6 y 13.3 se presentan algunos resultados teóricos relacionados con el problema del "unus mundus". Además, se introduce, en notas a pie de página, bibliografía adicional junto con referencias a puntos de vista no presentados en la primera edición del libro.

Dedico esta edición a la memoria de mis padres, Roberto Suárez Saitone y Blanca Antola Álvarez. También la dedico a la memoria de Eduardo Calcagno, en cuya biblioteca estudié por primera vez algunos de los temas de física, de matemática y de filosofía relacionados con la teoría de la relatividad.

Roberto Suárez Ántola
Montevideo, 31 de julio de 2019

Prólogo a la primera edición

Este libro tuvo su origen en tres conferencias sobre teoría de la relatividad y cosmología que el autor presentó, desde un punto de vista fisicomatemático y epistemológico, en la Universidad Católica del Uruguay.

Está destinado a un público interesado en los problemas que se plantean en torno al conocimiento científico en la denominada época postmoderna, y a la vez dispuesto a estudiar con cierto detenimiento y profundidad las bases fisicomatemáticas de la teoría de la relatividad, así como algunas de sus aplicaciones en astrofísica y cosmología.

Se supone que el lector posee conocimientos básicos de álgebra, cálculo infinitesimal de una y varias variables, y física general (elementos de mecánica, electromagnetismo y óptica, física térmica y cuántica), como los que pueden adquirirse durante los primeros dos años de estudios universitarios de ciencia o ingeniería. Teniendo esto último en cuenta, se presentan algunos temas complementarios de matemática y de física con los que el lector pudiera no estar familiarizado.

Sobre esta base se desarrollan los fundamentos de las teorías restringida y generalizada de la relatividad. Se consideran los resultados de experimentos y observaciones confirmatorios. Se señalan algunas dificultades de la teoría y se mencionan caminos alternativos que han sido o están siendo explorados.

No se supone que el lector esté familiarizado con la epistemología ni con otras ramas de la filosofía. Los contenidos filosóficos, en particular epistemológicos, se introducen a medida que se necesitan, comenzando con el primer capítulo que está destinado por completo a ellos.

Los capítulos dos, tres y cuatro describen el desarrollo de las ideas en la mecánica newtoniana de partículas y la teoría clásica de campos, incluyendo el electromagnetismo, empleando el principio de relatividad de Galileo como hilo conductor.

En el capítulo cinco se examina la crisis en los fundamentos de la física del siglo XIX que condujo a la teoría restringida de la relatividad.

Los capítulos seis, siete y ocho se destinan a estudiar los fundamentos, las limitaciones y algunas aplicaciones de la teoría restringida. Al final del capítulo ocho se examinan las relaciones epistemológicas entre las teorías de la física clásica luego de ser aceptada la teoría restringida de la relatividad.

El capítulo nueve se destina a complementar los conocimientos de matemática.

En particular, en este capítulo se exponen elementos de geometría diferencial de superficies y su extensión a variedades en n dimensiones, incluyendo nociones de cálculo tensorial, transporte paralelo, geodésicas y curvatura de variedades.

El capítulo diez se destina a exponer los fundamentos de la teoría generalizada de la relatividad. Las relaciones epistemológicas entre las teorías de la física clásica luego de ser aceptada la teoría generalizada de la relatividad se consideran al final del capítulo.

El capítulo once presenta algunas aplicaciones de la solución de Schwartzschild de las ecuaciones del campo gravitatorio y resume la evidencia confirmatoria disponible en relación con la teoría generalizada. Además, se introducen los fundamentos del cálculo de variaciones, cuyos resultados se emplean varias veces a lo largo de la exposición.

En los capítulos doce y trece se desarrollan a continuación, con cierto detalle, aplicaciones astronómicas y temas de cosmología física relativista, acompañados de consideraciones históricas y epistemológicas.

En el capítulo catorce se examinan la estructura de la teoría restringida de la relatividad, la teoría generalizada y la cosmología física desde el punto de vista epistemológico.

En el capítulo quince se consideran brevemente algunos aspectos del trabajo de investigación científica desde los puntos de vista psicológico y sociológico.

Se enfatizan los temas que permiten comprender los avatares en el desarrollo de las ideas científicas, hasta que desaparecen definitivamente en el olvido o pasan a formar parte de un paradigma estabilizado y dominante.

En el capítulo dieciséis se discute el principio de relatividad y algunas de sus consecuencias desde una perspectiva filosófica más amplia, enfatizando las limitaciones autoimpuestas de la ciencia.

Las referencias bibliográficas detalladas y algunos complementos importantes aparecen en forma de notas a pie de página.

Al final se sugiere una lista de libros que pueden ser de interés para reafirmar algunos conceptos o profundizar en los temas de ciencias fisicomatemáticas, cosmología y filosofía aquí presentados.

Se ha buscado, hasta donde ha resultado posible, que la exposición sea autocontenida, que las ideas físicas y filosóficas, así como las hipótesis básicas, se encuentren destacadas y explicadas, y que sus consecuencias se desprendan naturalmente durante el desarrollo de los temas.

Asimismo, se ha intentado hacer participar al lector en la construcción de la teoría, evitando por igual dos extremos. Uno de ellos consiste en fundamentar la exposición en herramientas matemáticas avanzadas, que por lo general el lector desconoce, pero que no se le explican, y que por tanto se ve obligado a aceptar o a estudiar en otra parte.

El otro extremo evita la necesaria fundamentación matemática recurriendo a enunciados puramente verbales y a representaciones gráficas, sugerentes pero muchas veces próximos a una fantasía sistematizada.

En la medida en que se alcancen estos objetivos, el libro debería presentar utilidad como material de estudio introductorio y como base para poder acceder a tratamientos más avanzados, en libros y en artículos de investigación, algunos de los cuales integran las referencias bibliográficas a pie de página.

Agradezco muy especialmente a los Profesores Mario Guerra y Carlos Zamalvide, quienes tomaron a su cargo una primera lectura crítica del texto. Sus sugerencias al respecto resultaron muy valiosas.

Asimismo, agradezco el estímulo a mi trabajo brindado por el director del Departamento de Formación Humanística, Padre Antonio Ocaña S.J.

Finalmente, este libro no hubiera existido de no ser por el decisivo apoyo de mi esposa.

Roberto Suárez Ántola
Montevideo, 31 de julio de 2011

1. Introducción

"Las dos suposiciones fundamentales de la filosofía natural son:

1. *Hay un mundo real externo que existe en forma independiente de nuestro acto de conocimiento.*
2. *El mundo real externo no es directamente cognoscible."*

Max Planck, "Where is Science going", W. W. Norton & Co., New York, 1932, p. 82.

"Creo que la epistemología es importante no solo para las ciencias individuales sino también para la filosofía, y que el malestar religioso y filosófico de nuestro tiempo, que a todos nos atañe, es sin duda en buena parte un malestar sobre la filosofía del conocimiento humano. Nietzche lo denominó nihilismo europeo, y Brenda la traición de los intelectuales. Yo desearía caracterizarlo como una consecuencia del descubrimiento socrático de que no sabemos nada; es decir, de que nunca podemos justificar racionalmente nuestras teorías.
Pero este importante descubrimiento, que ha generado, entre muchos otros malestares, el malestar del existencialismo, solo es un descubrimiento a medias; y el nihilismo puede superarse. Pues aunque no podamos justificar racionalmente nuestras teorías y ni siquiera podamos probar que son probables, podemos criticarlas en forma racional. Y podemos distinguir entre teorías mejores y peores.
Pero esto ya lo sabía, antes incluso que Sócrates, Jenófones, que dejó dicho:
***Los dioses no desvelaron todo a los mortales desde el principio; pero en el curso del tiempo podemos aprender y conocer mejor las cosas**."*
Karl Popper, en el prefacio de "En busca de un mundo mejor", Paidós, Barcelona, 2ª edición, 1994.

Durante nuestra primera infancia elaboramos el concepto de realidad objetiva: la existencia de un mundo externo, independiente del observador.
Los trabajos pioneros de Jean Piaget en psicología, el desarrollo de la neurociencia cognitiva, y la perspectiva evolutiva (darwiniana) de etólogos como Konrad Lorenz condujeron en la segunda mitad del siglo XX a un replanteo y a una superación del tratamiento filosófico previo de las categorías del conocimiento: objeto, espacio, tiempo, causalidad, número (cardinal y ordinal), cantidad (física).
Para ese entonces el desarrollo histórico de las ciencias de la naturaleza había contribuido a consolidar y a codificar el concepto de realidad objetiva en leyes, a menudo cuantitativas, como lo son desde hace siglos las leyes físicas.
Para que una relación numérica sea la expresión de una ley, los símbolos que toman valores numéricos deben presuponer la medición de propiedades de una cosa en un contexto, real o potencialmente experimentado por un observador.
Los contenidos de las ciencias son objetivos en la medida en que son reproducibles por cada observador y resultan iguales para diferentes observadores.
La objetividad de la ciencia, en este sentido, implica disponer de expresiones simbólicas para expresar y comparar los resultados de actos particulares intersubjetivos.
Cuando, como en la física macroscópica, se abstraen unos pocos elementos de la situación observacional o del contexto experimental, y los efectos de la interacción entre el observador y el objeto observado parecen no afectar a las propiedades de éste último, efectuar una distinción tajante entre sujeto y objeto, entre observador y observado, resulta plausible e inclusive conveniente. Esto es lo que ocurre tanto en la mecánica newtoniana como en la mecánica relativista.

Por el contrario, a partir de la aparición de la mecánica cuántica, la mencionada distinción, tal como se ha desarrollado en la física clásica, es cuestionada. No obstante, la validez de una distinción neta entre sujeto y objeto había sido cuestionada (mucho antes) en las ciencias humanas.

1.1 Mediciones de longitudes y tiempos

En disciplinas como la Física, la Química, la Astronomía y la Fisiología es preciso trabajar con propiedades que cuando se miden originan cantidades que se consideran continuas. La posibilidad de medir cantidades continuas, como longitudes, ángulos, tiempos, y masas, requiere que sea factible:

(a) Definir una cantidad conservada mediante operaciones con la materia: la masa de un trozo de arcilla plástica que puede adoptar diferentes formas, un volumen de agua que se puede verter en recipientes de diferentes formas, o la longitud de una regla que se puede ubicar en distintos sitios y con distintas orientaciones.
(b) Combinar objetos separados de tal forma que una cantidad conservada asignable a la combinación sea la suma de las cantidades correspondientes a las partes tomadas por separado.
(c) Introducir una unidad de medida, sus múltiplos y submúltiplos, para poder traducir cantidades en números[1].

La elección de cuerpos para la medición de longitudes depende de la percepción de las relaciones entre los cuerpos materiales y sus modificaciones. La percepción visual del entorno sugiere que algunos cuerpos se deforman o fluyen en forma muy evidente, mientras que otros parecen permanecer incambiados durante meses o años. Entre estos últimos los hay alargados, como las varas de madera, que pueden emplearse a modo de unidad para comprobar la constancia de las dimensiones de los cuerpos aparentemente estables.
La elección de tal o cual regla para medir longitudes es en buena medida arbitraria. Se justifica en última instancia por la simplicidad y constancia de las relaciones métricas entre los cuerpos que se descubran mediante su empleo, así como por la facilidad con que la conexión entre esas relaciones métricas y otras cantidades físicas permitan inferir regularidades en el comportamiento de los cuerpos.

Algo similar ocurre con las medidas de tiempo[2]. En principio cualquier proceso periódico puede emplearse como reloj: por ejemplo el latido de un corazón. Pero algunos procesos periódicos, como la alternancia entre el día y la noche asociada a la rotación de la Tierra, cuando se emplean como relojes permiten descubrir relaciones cronológicas simples, constantes y significativas en un conjunto muy extenso de

[1] Se supone que estos números son números reales cualesquiera.

[2] Sobre el concepto de tiempo y su medida puede consultarse los artículos de: A. Danjon "Le temps. Sa définition pratique, sa mesure", L'Astronomie, vol. xliii, 1929, pp. 13-22; G. Clemence "Time and its measurement", Astronomical Society of the Pacific Leaflets, Vol. 6, 1952, pp. 259-266; J. Derral Mulholland "Measures of time in astronomy" Publications of the Astronomical Society of the Pacific, vol. 84, Nº 499, 1972, pp. 357-364; y P. Aichelburg "On the evolution of the concept of time and its implication for modern cosmology", en Science and Religion: Global Perspectives, Philadelphia, PA, USA, 2005.

procesos naturales terrestres y celestes: la elección de un reloj se justifica, como la elección de una regla para medir longitudes, por su utilidad tanto teórica como práctica.

A medida que se perfecciona el conocimiento del mundo físico, la elección de las unidades de medida y los procedimientos de medición también se perfeccionan. A su vez, el empleo de las nuevas unidades y procedimientos de medición permiten poner en evidencia regularidades más sutiles que hasta ese momento habían pasado desapercibidas y plantea nuevos problemas a la física teórica.

1.2 Geometrías

Los cuerpos materiales, macroscópicos, pueden en principio verse y tocarse: se conciben como objetos que conservan su identidad bajo diferentes perspectivas sensoriales, incluso cuando no son percibidos por nuestros sentidos.
El estudio cuantitativo de los cuerpos materiales con ayuda de medidas de longitudes y de ángulos, conduce al concepto de cuerpo rígido (cuerpo en el cual las distancias entre todos sus puntos, tomados dos a dos, permanecen invariables). El estudio de las relaciones entre los cuerpos origina el concepto de espacio físico.
La posibilidad de dividir un cuerpo extenso una y otra vez sugirió la introducción de una idealización: el punto material o partícula. Desde este punto de vista el espacio físico se puede considerar como un sistema de relaciones entre cuerpos materiales compuestos por puntos.
Un modelo matemático de este espacio físico es lo que conocemos como una **geometría física,** para distinguirla del concepto de geometría tal como se lo concibe en la matemática actual: lo que puede denominarse con mayor propiedad **geometría matemática**. Los puntos, las formas y las estructuras de esta última no tienen por qué tener relación con el espacio físico: a menudo son abstracciones construidas abstrayendo a partir de puras abstracciones de un nivel inferior. Así los espacios de la geometría matemática son conjuntos de puntos cualesquiera entre los que se definen ciertas relaciones mediante postulados. Son, pues, espacios abstractos cuyos puntos constitutivos pueden ser, por ejemplo, funciones.
Desde el punto de vista matemático, una geometría consiste en el estudio de las propiedades de conjuntos de puntos de un espacio (figuras) que permanecen invariantes al transformar cada punto del espacio abstracto en otro, empleando para ello transformaciones de un determinado tipo (más precisamente, lo que en matemática se conoce, desde el siglo XIX, como grupo de transformaciones).

El modelo matemático de espacio físico no tiene que coincidir con el de Euclides, pese a que este último se corresponde mejor con el concepto de espacio que se desarrolla en forma progresiva y más o menos tardía como parte de las capacidades cognoscitivas que se adquieren durante la infancia y la adolescencia. Al parecer fue seleccionado durante la filogenia por su valor evolutivo para la especie, junto con los conceptos de objeto, de flujo del tiempo independiente de un espacio ilimitado, y de causalidad asociada al establecimiento de una conexión entre sucesos contiguos en el tiempo, uno de ellos, la causa, que aparece primero y otro, el efecto, que aparece después de o a lo sumo simultáneamente con la causa.

1.3 Ciencia y principio de causalidad

Este concepto de relación causal evolucionó hasta convertirse en el concepto de **causalidad eficiente** propio de los estudios científicos y filosóficos desde la época de la Grecia clásica hasta la actualidad.
Las personas parten de un mundo de sensaciones y percepciones, sentimientos y voliciones, habla interna y cualidades secundarias (qualia) que en forma natural se asocian con un mundo externo. Pronto detectan regularidades: el resplandor del alba da paso a la salida del sol, el destello de un relámpago es seguido por el retumbar del trueno. Encuentra resistencias y dificultades: mover un cuerpo masivo exige un esfuerzo, retener un líquido entre las manos en general no resulta posible. Pero en el fondo todo lo que percibimos son de algún modo solo apariencias, eventos simultáneos o sucesivos, cercanos entre sí o lejanos, asociados al mundo exterior o parte de vivencias íntimas.

Cuando se pone énfasis en los fenómenos, en las apariencias, la noción de causalidad pierde todo interés: desde el punto de vista meramente fenoménico la conexión entre el resplandor del alba y la salida del sol queda al mismo nivel que la conexión entre el relámpago y el trueno. Eso explica el rechazo de la causalidad como noción objetiva desde la crítica de David Hume en la primera mitad del siglo XVIII hasta las disquisiciones de filósofos contemporáneos como David Lewis.

Las conexiones causales interpretadas como regularidades permanentes son, efectivamente, solamente conjeturas.
No obstante, si admitimos que nuestras percepciones evocan cosas que se encuentran y acontecimientos que se producen en un mundo externo, existente por sí mismo y hasta cierto punto susceptible de ser conocido, el problema de la causalidad se puede plantear en otros términos, en los que el problema del ser se funde con el problema del conocer en una unidad indisoluble.
La relación causal entendida como conexión objetiva entre eventos que acontecen en un mundo existente por sí mismo, con independencia de posibles observadores, adquiere pleno sentido en un marco de conocimientos cuya construcción y modificación progresivas presupone a su vez la búsqueda de relaciones y explicaciones causales.
Que un relámpago es causa objetiva de un trueno es una conjetura que se confirma cuando se comprende que una descarga de electricidad produce una onda de choque en la atmósfera que se percibe como trueno, y en pequeña escala se reproduce lo esencial de la situación en un laboratorio. El trueno que se escucha asociado al vuelo de un avión supersónico, refuerza la validez de la conexión causal entre relámpago y trueno, debido a un fenómeno intermedio común a ambos casos: la onda de choque, que se produce en las alas del avión y se escucha luego de que éste ha pasado.
Por el contrario, el resplandor del alba se asocia con, pero aunque aparece siempre primero no es causa objetiva de, la salida del sol. Ambos son efectos de la rotación de la Tierra respecto de su eje, y quedan explicados por esa rotación.

En la medida en que se busquen explicaciones causales de los acontecimientos y no una mera descripción de los mismos, no es posible eludir los problemas filosóficos asociados con la causalidad, y en particular la discusión del principio de causalidad.
A veces se formula el principio de causalidad como la afirmación “todo efecto tiene una causa”. Esta formulación ha conducido a confusión al discutir el carácter lógico y epistemológico del principio. Un análisis cuidadoso muestra que ni es evidente en sí

mismo, ni se lo puede demostrar como proposición analítica. Es ni más ni menos que un supuesto necesario para todo conocimiento científico de la realidad.
Al respecto dice Messer[3]: "Para que los cambios nos resulten comprensibles, necesitamos referirnos a sus causas. Suponemos que todo cambio tiene su causa. Este principio es válido en su universalidad no por fundarse en la experiencia, pues entonces deberíamos haber probado su validez en todas las experiencias posibles. Pero por otra parte no debemos temer que pueda ser desmentido por la primera experiencia que se presente. Si no pudiéramos encontrar a un cambio ninguna causa no nos contentaríamos con pensar que no la tiene, sino que creeríamos que la causa no es provisionalmente desconocida. El principio es válido independientemente de la experiencia…
Pero si podemos considerar de este modo el principio de causalidad como una condición a priori de la experiencia, no por eso resulta válido a priori, exactamente en el mismo sentido en que lo son los principios de la matemática pura y de la lógica. No es lógicamente necesario, su negación no implica una contradicción. El concepto de cambio no contiene el concepto de causa, no contradecimos el contenido del primero si afirmamos que un cambio no tiene causa. Pero no podríamos obtener conocimiento científico alguno de un cambio semejante… Frente a él se nos pararía, por así decirlo, el intelecto. La afirmación de que todo lo existente haya de ser comprensible para nosotros no es una proposición lógicamente necesaria; es solo un supuesto y, por ende, el principio de causalidad solo tiene el valor epistemológico de un supuesto."

Para los fines de un estudio como éste, que enfoca el problema del conocimiento desde la perspectiva suministrada por un análisis crítico de las ciencias físicas, admitiremos el planteo de conexiones causales solamente entre eventos, es decir entre cambios de estado en las cosas o en partes de una misma cosa. Un evento C que acontece en una cosa o en parte de una cosa α, causa un evento E (efecto) en otra cosa o en otra parte de la misma cosa β, cuando y solo cuando C genera una transferencia de energía de α hacia β de la que resulta E. Las conexiones causales se reducen a cambios asociados en última instancia a flujos de energía en el mundo real[4]. Más adelante precisaremos el alcance y

---

[3] A. Messer, "*Einfuhrung in die erkenntnistheorie*", F. Meiner. Leipzig, 1909, p. 42.

[4] Parece conveniente señalar que los filósofos contemporáneos que continúan la tradición empirista, consideran que el problema de la causalidad se debe plantear solamente a nivel del lenguaje, en términos de enunciados conocidos como contra-fácticos.
Ese planteo se puede resumir así: "A y B son acontecimientos causalmente conectados de modo que A es causa de B siempre y cuando si A no hubiera acontecido entonces B tampoco hubiera acontecido".
Así planteada las cosas, la conexión regular observada entre el resplandor del alba y la salida del sol, interpretada como una regularidad permanente conjeturada, podría quedar comprendida dentro de este planteo siendo A el resplandor del alba y B la salida del sol. Pero como este planteo de la causalidad no insiste en el orden temporal de los acontecimientos, también se lo puede aplicar cuando A es la salida del sol y B es el resplandor del alba. Lo mismo puede decirse de la conexión entre el relámpago y el trueno. Pero mientras la conexión entre el resplandor del alba y la salida del sol no es una conexión causal en el sentido científico, la conexión entre el relámpago y el trueno sí lo es.
Desde un punto de vista lógico, un enunciado contra-fáctico no es propiamente una proposición, puesto que no se le puede asignar un valor de verdad. Esto no constituye un problema para las personas que propugnan este abordaje basado en enfatizar las regularidades aparentes. Pero además, la forma contra-fáctica del enunciado de la causalidad lo hace dependiente del enunciado "Si A acontece, entonces acontece B", que es el enunciado de una ley. Lo cual tampoco constituye un problema si la idea es permanecer a nivel de los fenómenos ignorando los noúmenos (cosas) del mundo externo. Sobre el principio de causalidad en el marco de una filosofía de sistemas ver M. Bunge "*Emergencia y Convergencia*", Gedisa, Barcelona, 2004.

las limitaciones de este punto de vista. Retomemos ahora el tema de esta introducción, orientada a revisar algunos conceptos básicos para el planteo del principio de relatividad.

1.4 Principios de relatividad y teorías de la relatividad

Las posiciones relativas de los cuerpos a menudo cambian. Para cuantificar esta variación se emplean **marcos de referencia**. Un marco de referencia está dotado con elementos para medir longitudes (reglas) e instantes de tiempo (relojes) y se forma a partir de un sistema de cuerpos en reposo relativo.
Desde el punto de vista físico la descripción de los cambios en un espacio cuyos elementos son puntos se hace mediante eventos puntuales. Un **evento puntual es algo que acontece en un instante en algún punto del espacio.** Para distinguir un evento puntual de otro se debe medir o especificar el instante de tiempo en el que se produce y las coordenadas de posición del lugar donde se produce, respecto de un mismo marco de referencia. Un mismo cuerpo puede cambiar de posición entre dos instantes de tiempo dados. El cambio de posición se representa mediante un desplazamiento. La velocidad del cuerpo se define entonces como desplazamiento por unidad de tiempo.
Un modelo matemático de estos cambios, que permita describirlos utilizando medidas de posición y de tiempo, relativas a un sistema de cuerpos materiales que se toman como base para construir un marco de referencia, es una **cinemática**.

La descripción del movimiento presupone tratar con objetos que, directa o indirectamente, pueden ser observados y medidos. Pero esto presupone a su vez que las cosas estudiadas interactúan con su ambiente. Para una descripción más completa del objeto, esa interacción debe producirse en diferentes instancias y bajo diferentes condiciones.
La introducción del concepto de fuerza de interacción entre cuerpos materiales y de leyes del movimiento, en un modelo matemático que involucra esas fuerzas entre los cuerpos y permite predecir sus movimientos, conduce a una **dinámica**.
Si se admite que un mismo fenómeno se debe poder describir respecto de diferentes marcos de referencia, o en forma equivalente se debe poder describir por distintos observadores, empleando las mismas leyes dinámicas, se introduce un **principio de relatividad**. Con su ayuda se pueden relacionar entre sí las descripciones temporales y espaciales correspondientes al mismo suceso en diferentes marcos de referencia.

En Física se denomina **teoría de la relatividad** a toda teoría cuyo objeto es la discusión de las elecciones de marcos de referencia respecto de los cuales se formulan las leyes de la naturaleza de forma tal que un hecho objetivo predicho respecto de uno de esos marcos de referencia, tal como una colisión entre partículas, resulte predicho en cualquier otro marco de referencia admisible.
Para ello, la forma de las ecuaciones que expresan leyes naturales debe permanecer invariante al cambiar de un marco de referencia a otro, dentro de la clase de marcos considerada. Pues si la forma de las ecuaciones se modificara al pasar de un marco a otro, podría ocurrir que un suceso objetivo predicho respecto de un marco no resultara predicho respecto de otro marco de su misma clase.
Desde este punto de vista la mecánica clásica comprende también un principio de relatividad, denominado principio de relatividad de Galileo, y una teoría de la relatividad basada en este principio. Pero el aspecto relativista de la mecánica de Newton permaneció siempre en un segundo plano.

Cuando se hace referencia a una teoría de la relatividad, se asume en forma implícita que se trata de **la teoría de la relatividad** basada en el denominado **principio de relatividad de Einstein**, puesto que en esta teoría el punto de vista relativista pasa a un primer plano.

La introducción de un principio de relatividad por lo general restringe la clase admisible de marcos de referencia que pueden utilizarse para la descripción del movimiento de los cuerpos y sus interacciones.
En el caso de la mecánica newtoniana y también en el caso de la teoría restringida o teoría especial de la relatividad, se trabaja con sistemas de referencia denominados inerciales. En la teoría generalizada de la relatividad esta restricción se suprime.

Uno de los resultados principales de la teoría de la relatividad, tanto en su versión restringida como en su versión generalizada, es la interrelación entre espacio y tiempo. Esta interrelación se manifiesta cuando se conectan entre sí las posiciones y los instantes de un mismo evento puntual respecto de dos marcos de referencia distintos.
Se la puede representar mediante un espacio abstracto cuyos puntos representan eventos. Cada marco de referencia define un sistema de coordenadas de tiempo y espacio que permiten caracterizar los eventos mediante cuaternas de números reales.
La cinemática, en tanto mera descripción del movimiento, se reduce a una geometría en un espacio no euclidiano tetra-dimensional.
La dinámica, en cuanto descripción del movimiento atendiendo a sus causas, se reduce a la identificación de curvas asociadas con las partículas y con sus interacciones. Aparece entonces una nueva geometría física, mucho más abstracta que la geometría física propia de la física newtoniana.
Mientras que en esta última geometría física los objetos geométricos están constituidos por conjuntos de puntos considerados en un mismo instante de tiempo, en la geometría propia de la relatividad los objetos geométricos están constituidos por conjuntos de eventos. Entonces, como veremos posteriormente, respecto de un mismo marco de referencia los eventos de una misma figura pueden tener coordenadas temporales diferentes.

1.5 Ciencia, filosofía y matemática

La teoría especial de la relatividad se emplea en el diseño de aceleradores de partículas para aplicaciones médicas, industriales y de investigación científica.
Permite comprender el origen de y calcular la energía liberada por la fisión de núcleos atómicos tanto en los reactores nucleares (utilizados para generar energía eléctrica, para producir isótopos para uso médico e industrial y para la investigación) como en las armas atómicas de destrucción masiva.

La teoría generalizada de la relatividad suministra a su vez la base física y matemática para la mayor parte de las cosmologías actuales, en particular para las cosmologías más conocidas y aceptadas por los científicos contemporáneos.
Con la teoría restringida el tiempo, habiendo perdido su carácter absoluto, pasa a depender del movimiento. Con la teoría generalizada pasa a depender, además, de la distribución de la materia a través de los efectos gravitatorios. Tanto la influencia del movimiento como la influencia de la gravedad se deben tener en cuenta para sincronizar los relojes que se hallan ubicados en satélites, a distancias del orden de los 20 000 km

de la superficie terrestre, y que en conjunto hacen posible el funcionamiento del Sistema de Posicionamiento Global (GPS).
Esto sugiere la importancia (tanto teórica como práctica) de la física relativista y el interés que presenta su estudio.

Pero además, existe un motivo más profundo, de índole filosófica. Todo desarrollo científico de gran envergadura trasciende la mera búsqueda de explicación y control de los acontecimientos en su área específica, para incidir sobre la visión del mundo de la época, modificándola[5]. Es ciencia que engendra meta-ciencia, o si se quiere "filosofía natural" en el sentido ahora algo anticuado con el cual Max Planck usa el término en la cita que da comienzo a este capítulo.

Las concepciones del mundo propiamente dichas trascienden a la ciencia e involucran aspectos que generalmente se consideran en el dominio de la filosofía.
Llegados a este punto, conviene distinguir entre filosofía especulativa y filosofía criticista, como lo hace Broad[6].

La filosofía criticista tiene como fin el analizar el significado de y las relaciones entre los conceptos básicos de la ciencia, así como detectar y explicitar las suposiciones y creencias utilizadas tanto en la vida cotidiana como en el quehacer científico y en la filosofía especulativa.
Suministra un marco conceptual para analizar temas tales como la aplicabilidad de la matemática a las ciencias físicas, el origen y el rol de las leyes matemáticas de conservación, y la introducción y la evolución del concepto de sustancia.
Uno de los aportes de la filosofía criticista, de particular importancia en relación con el principio de la relatividad y su evolución histórica, es el denominado principio de identidad.
Este principio resume la tendencia a priori de la mente humana a negar la diversidad, buscando la identidad en una realidad que, por así decirlo, se le resiste[7]. La medida en que la unificación es o no posible depende de sucesivos intentos, del ensayo y el error: solo a posteriori se puede saber qué aspectos de la realidad ceden a esta tendencia.
El principio de identidad cristaliza en dos niveles.
En un nivel que puede denominarse de segundo orden, a través de la búsqueda de leyes que se supone se aplican en todo lugar y en todo tiempo. Las predicciones que estas leyes permiten efectuar poseen valor evolutivo para la especie humana.
En un nivel que se puede denominar de primer orden, el principio de identidad se manifiesta a través de la búsqueda de explicaciones causales y leyes de conservación. Esta búsqueda presupone la posibilidad de hallar, en las cosas concretas, sustancias que

[5] La importancia de la revolución copernicana no estuvo en suministrar la posibilidad de un cálculo más exacto de los movimientos de los planetas en comparación con la teoría geocéntrica de Ptolomeo (cosa que por otra parte no hacía) sino en haber arrancado al hombre con su micro-cosmos y su mundo sub-lunar del centro de un universo acotado para arrojarlos a un espacio infinito. En eso consistió el cambio más radical de la visión del mundo acaecido durante el siglo XVI, cambio que a su vez hizo posible la ciencia moderna.

[6] C. Broad "*El pensamiento científico*", introducción, Tecnos, Madrid, 1963.

[7] E. Meyerson "Identidad y realidad", Reus, Madrid, 1921.

perduran a través de todos los cambios[8]. Un ejemplo es la energía, que se asume es transferida cuando hay una conexión causal entre eventos, como se estableció en 1.3 a propósito de la discusión sobre el principio de causalidad. Otro ejemplo es la masa tal como se la concebía en la física y química clásicas, previamente a la aparición de la teoría de la relatividad.

La filosofía especulativa, en la medida en que pretende abarcar el mundo como totalidad y dar una respuesta a la pregunta sobre el puesto del hombre en el cosmos, refleja las esperanzas, temores y aversiones del filósofo. Aun así, los aportes de la filosofía especulativa son importantes desde el punto de vista científico, sobre todo a través de sus métodos y de sus fracasos. Sus métodos porque ayudan a evitar la sobre-simplificación en la que cae fácilmente el especialista, sus fracasos porque recuerdan la complejidad del mundo y alertan acerca de los peligros de extrapolar sin fundamento los contenidos y los métodos de la ciencia.
Como la transición de la filosofía especulativa a la filosofía criticista, y de esta última a las ciencias, se produce sin solución de continuidad, ignorar lo que se ha hecho en filosofía especulativa puede ser, además de inconveniente, riesgoso[9].

Puesto que la filosofía abarca los intentos de la humanidad de construir concepciones del universo reflexionando sobre su conducta valorativa teórica y práctica, el término filosofía natural[10], algo más restringido, parece adecuado para aludir a la ciencia cuando se la considera como productora de meta-ciencia.
En la medida en que la filosofía implique una reflexión sobre la conducta valorativa teórica, incluye a la teoría de la ciencia, tanto formal (lógica: reflexión sobre pensamiento correcto) como material (teoría del conocimiento: reflexión sobre el pensamiento verdadero, sobre la posibilidad y el origen del conocimiento, y sobre las categorías del conocimiento).

La primer suposición de la filosofía natural que señala Planck, la existencia de una realidad externa al sujeto cognoscente, no parece ni demostrable ni refutable por medios puramente racionales. Debe admitirse como un postulado, que en todo caso encuentra un apoyo irracional en una vivencia inmediata de la realidad a través de las resistencias que las personas siempre encuentran frente a su querer y desear.
La segunda suposición, la ausencia de conocimiento directo de esa realidad admitida, se refiere al mundo impersonal de los objetos reales y sus relaciones, que se encuentran detrás de nuestras sensaciones.
Tomada como un postulado, justifica la tarea auto-impuesta por los científicos, de salir (por así decirlo) a la caza de la realidad apoyados en un método riguroso (experimental-hipotético-deductivo-experimental) y en un cuerpo de conocimientos previos que se va modificando sobre la marcha a medida que aparecen resultados de experimentos y

---

[8] Ver E. Zahar "Meyerson and the role of mathematics in physical discovery" The British Journal of the Philosophy of Science, vol 31, Nº 1, 1980, pp. 1-43.

[9] Esto último se observa especialmente en algunos de los esfuerzos tendientes a divulgar conocimientos sobre el origen y evolución del cosmos empleando formulaciones verbales del modelo cosmológico estándar, incluyendo aspectos relacionados con la gravedad cuántica.

[10] Que por otra parte aparece en el título de la obra más importante y conocida de Newton: Principios matemáticos de la filosofía natural.

observaciones que no pueden ser explicados, o se descubren inconsistencias que habían pasado desapercibidas entre cuerpos de teoría establecidos que por lo menos en parte se refieren a los mismos aspectos de los fenómenos estudiados[11].

La caracterización del estado de algunos objetos mediante variables numéricas medibles y la búsqueda de relaciones cuantitativas entre variables introdujo los modelos matemáticos en las ciencias de la naturaleza, e hizo posible la aparición y el desarrollo de la física en su acepción moderna: tener un fundamento experimental y estar expresada de forma matemática.
Por su parte la física estimuló la aparición de nueva matemática, más abstracta y de mayor alcance. Gracias a la nueva matemática, la física relativista y cuántica pudo ser formulada con precisión y sometida a confirmación experimental.
Esto permitió modificar tanto nuestra comprensión como nuestra visión del mundo en un nivel sin precedentes.
El camino que condujo a la física-matemática contemporánea comenzó en Grecia con la geometría sintética de Euclides y la acústica musical de Pitágoras.
La tradición oral sostiene que Pitágoras pasaba a menudo por un taller de herrería, en su isla natal de Samos. Parece que se detenía a observar cómo el herrero martillaba sus piezas metálicas. Un día se dio cuenta de que al ser percutidas, barras de longitudes diferentes producían sonidos diferentes, y que la altura del sonido guardaba relación con las dimensiones del objeto martillado. Cuando volvió a su casa comenzó a experimentar con las cuerdas de una lira. Descubrió que dividiendo una cuerda por uno de los números naturales 1, 2, 3, 4 cuya suma es 10, y haciéndola vibrar, se producía lo que en música se conoce como armonía: una octava, una quinta y una cuarta.
Desde el comienzo de la edad del hierro los herreros martillaban y los transeúntes escuchaban los mismos sonidos que escuchó Pitágoras. Para la mayor parte de ellos se trataba de un ruido, inclusive un ruido molesto. Para Pitágoras, ese día feliz, el ruido se transformó en información[12]. Como dijo una vez Albert Szent-Györgyi: "El descubrimiento consiste en ver lo que todos han visto y pensar lo que nadie ha pensado".

La geometría sintética de Euclides suministró la primera formulación de un sistema axiomático, que apelaba a la intuición y se apoyaba en proposiciones cuya verdad parecía evidente. El ejemplo más famoso es el axioma de las paralelas, cuya discusión y sustitución condujo a las geometrías no euclidianas.
La geometría euclidiana en su formulación original era una teoría matemática en sí[13], pero también una teoría de los cuerpos materiales y sus relaciones intemporales, es decir

---

[11] Desde una concepción filosófica instrumentalista, se considera que los científicos no buscan la verdad: se conforman con la construcción de modelos más o menos exitosos. Para los que adoptan esta postura pragmática carece de sentido decir que un modelo es más verdadero que otro: un modelo apenas puede ser más exitoso que otro, midiendo el éxito a través del balance explicativo, es decir la capacidad de explicar y predecir fenómenos observables. La pretensión de salir a la caza de la realidad, en el sentido de emprender la búsqueda de la verdad, carece entonces de sentido para los instrumentalistas.

[12] Arthur Koestler *"The act of creation"*, Hutchinson & Co, New York, 1964.

[13] El matemático inglés Hardy decía, a principios del siglo XX, que al leer a Euclides le parecía estar en presencia de un miembro de otra universidad ("a fellow of another college").

una geometría física. El punto, la línea y las figuras planas y espaciales se introducían en correspondencia con e idealizando a cuerpos materiales.

El descubrimiento que los números racionales no bastaban para expresar todas las relaciones métricas de la geometría euclidiana, atribuido a Pitágoras[14], condujo a la ampliación del conjunto de los números racionales al conjunto de los números reales. La resolución de ecuaciones algebraicas llevó en el siglo XVI a ampliar el conjunto de números reales introduciendo los números complejos, que resultaron fundamentales para formular matemáticamente la física cuántica en el siglo XX.
La formalización de la mecánica clásica se acompañó del desarrollo del cálculo infinitesimal y de la geometría diferencial de curvas y superficies en el espacio euclidiano, a partir del siglo XVII. También en el siglo XVII aparece la innovación más importante en geometría, desde la época de Euclides: Descartes sienta las bases de la geometría analítica.
De ahí en más geometría y algoritmos algebraicos primero, y geometría y cálculo infinitesimal después, se desarrollan interrelacionados.
Profundizando en la vía abierta por Descartes, Gauss demostró la posibilidad de construir una geometría intrínseca a una superficie, utilizando las curvas geodésicas, curvas de longitud mínima que conectan dos puntos de la superficie, están contenidas en ella y cumplen el rol de las rectas de la geometría euclidiana. Esto permite trabajar en forma independiente del espacio en el que se encuentre inmersa la superficie y la geometría resultante es común a todas las superficies que se pueden obtener a partir de la primera deformándola sin modificar las distancias entre puntos definidas por las longitudes de arco de geodésica que conectan cada par de puntos entre sí.

La fundamentación del cálculo en el siglo XIX, motivada por la vaguedad y algunos errores que acompañaron su desarrollo durante el siglo XVIII, dio origen al análisis matemático moderno, así como a una definición y teoría rigurosa del continuo.
Esto a su vez hizo posible la generalización de la geometría diferencial a una teoría de variedades diferenciables de n dimensiones, cuyos puntos venían representados por conjuntos ordenados de números reales entre los cuales se introducían un concepto de distancia no euclidiano. Riemann extendió a este caso más general el concepto de geometría intrínseca (incluyendo la generalización del concepto de curvatura intrínseca debido a Gauss). La investigación de geometrías abstractas definibles en dichas variedades y el desarrollo del análisis tensorial como la herramienta analítica más adecuada para esa investigación, preparó el terreno e hizo posible la formulación de la teoría general de la relatividad.

A fines del siglo XIX, tanto la mecánica newtoniana como la electrodinámica se habían organizado como sistemas matemáticos formales acompañados de reglas de correspondencia entre algunos conceptos matemáticos y los datos de la experiencia.

[14] No se conocen escritos de Pitágoras. Fundó una comunidad religiosa y política en Crotona, Italia, a donde emigró desde su Samos nativa cuando contaba con 40 años de edad. Al parecer se le ocurrió la idea de que los números naturales son la fuente de todas las cosas. Dicen que cuando se dio cuenta de que había demostrado que la diagonal de un cuadrado no se relaciona con su lado mediante un cociente de números naturales, se deprimió tan profundamente que consideró que no valía la pena continuar con vida. Resulta curioso que en la antigüedad se hacía referencia a los pitagóricos y se solía evitar nombrar a Pitágoras mismo.

Ambas presentaban un grado muy elevado de confirmación experimental pero eran inconsistentes entre sí.
En general, en física teórica, cuando dos sistemas formales se encuentran en una situación semejante, cabe esperar que las inconsistencias se deban a suposiciones no conectadas directamente con las mediciones o las operaciones físicas en el laboratorio. Si es así, se pueden eliminar esas suposiciones no operacionales sin perder el grado de acuerdo con la experiencia ya alcanzado. Más aún, ese grado algunas veces se puede ampliar significativamente.
Eso es lo que ocurrió cuando Einstein se dio cuenta de que el concepto de simultaneidad absoluta de eventos[15], propio de la mecánica newtoniana, no tenía fundamento operacional alguno porque no se había establecido un procedimiento para sincronizar relojes situados en puntos diferentes. Como consecuencia, aunque psicológicamente era (y sigue siendo) muy natural, el concepto absoluto de simultaneidad se puede eliminar de los fundamentos de la cinemática sin daño alguno para el resto del sistema. El motivo de fondo para abandonarlo es que si se mantiene, impide la integración de la mecánica con la electrodinámica en un todo coherente.

El insistir en la definición operacional de los conceptos físicos es importante, pero es un medio destinado a un fin: eliminar suposiciones infundadas de un cuerpo de teoría física. De ninguna manera puede considerarse como fin en sí mismo.
El siguiente diálogo relatado por Philipp Franck resulta muy ilustrativo[16]:
"*Philipp Franck a Einstein*: ¿Cambió usted de opinión, o cree usted aún que cada concepto en el discurso científico debería tener un significado operacional?
Einstein a Philipp Franck: Vea usted, yo no diría eso. Yo no diría que cada palabra debe tener un significado operacional, pero sí que se requeriría que en las consecuencias de cada palabra utilizada debe haber un significado operacional."

El énfasis en las definiciones operacionales de los conceptos encuentra un fundamento filosófico en la escuela del Empirismo Lógico de la primera mitad del siglo XX, heredera de la Filosofía Positivista del siglo XIX.
La filosofía positivista, de la cual Comte es uno de los exponentes más representativos, dio una nueva forma a la filosofía natural que emergió en el siglo XVIII a partir del éxito de la mecánica newtoniana y sustituyó a los remanentes de filosofía escolástica.
Los filósofos positivistas, al extender el marco conceptual de la nueva mecánica más allá de su ámbito natural, desestimaron parte de los problemas filosóficos tradicionales considerándolo carentes de sentido y pretendieron hacer una filosofía que fuera científica.
Cuando, a fines de siglo, el esquema mecanicista entró en crisis, y surgieron las teorías de la relatividad y de los cuantos para resolverla, resultó que algunas preguntas que parecían tener sentido en el marco del mecanicismo, ahora no lo tenían.
La insistencia de los empiristas lógicos en adherirse estrictamente a la experiencia y desconfiar del sentido de las conceptualizaciones previas, fue una primera respuesta adecuada a la necesidad de cambiar de paradigma. Pero luego cristalizó en una filosofía

[15] Es decir, si dos sucesos ocurren en el mismo instante para un observador, también son simultáneos para cualquier otro observador, independientemente de su estado de movimiento respecto del primero.

[16] Philipp Franck hizo referencia al diálogo con Einstein en el Coloquio sobre Filosofía de la Ciencia realizado en Boston, Massachussetts, en diciembre de 1961.

todavía más restrictiva que la del positivismo. Como la gente siempre se plantea los mismos problemas fundamentales pese a la opinión de los filósofos, la expansión de la perspectiva propia del Empirismo Lógico, sobre todo en los países de habla inglesa, dejó el camino libre a la filosofía existencialista en el resto[17] y contribuyó al malestar filosófico y religioso actual al que alude Karl Popper en la segunda cita al comienzo de este capítulo.

Puesto que las categorías que hacen posible nuestra experiencia y cognición dependen de nuestras características psicofísicas y de los sistemas simbólicos que nos suministra la cultura a la que pertenecemos, en este sentido se puede afirmar sin duda alguna que nuestro conocimiento es relativo. Pero de esto no se desprende necesariamente que todo conocimiento humano se limite a lo convencional y lo utilitario, o inclusive que en última instancia todo conocimiento humano sea fútil, como ha venido sosteniendo el relativismo epistemológico.
Sea cual fuere el significado metafísico de la realidad, puesto que los seres humanos son actores en el mundo y no meros espectadores, parecería razonable suponer que el aparato psicofísico debe guardar un tipo de correspondencia con esa realidad. Esa correspondencia debe ser suficiente para permitir la supervivencia de la especie en circunstancias cambiantes: nuestro aparato psicofísico no puede estar completamente desacoplado de las cosas, o acoplado en forma fortuita. Debemos reconocer con Üexkull que el mundo (umwelt) es en cada especie biológica un producto del correspondiente sistema de recepción, procesamiento y respuesta: hay un mundo para la mosca, otro para el erizo de mar, y otro mundo para el homo sapiens. No obstante, en este último caso hay una diferencia notable, que Bertalanffy señaló ya en 1937 [18] : a través del desarrollo del conocimiento científico se produce una progresiva des-antropomorfización que elimina las características específicas de la experiencia del homo sapiens. Ese proceso avanza a lo largo de tres direcciones[19].
En una primera dirección la invención de instrumentos de medición produce la expansión de lo observable más allá de las posibilidades de los sentidos y la reflexión sobre los hallazgos experimentales obtenidos se acompaña de la fusión de disciplinas directamente vinculadas con la experiencia sensorial, dando origen a nuevas disciplinas, más abstractas, que no guardan relación directa con los sentidos.
En una segunda dirección las constantes físicas se miden cada vez con mayor precisión con métodos independientes y los resultados convergen, sugiriendo que esas constantes representan aspectos objetivos de la realidad, más allá de sesgos biológicos o culturales.
En una tercera dirección, desaparece progresivamente todo lo que aparenta ser demasiado humano y lo que queda es un sistema de relaciones matemáticas[20]. Las teorías físicas, cada vez más, poseen contenidos que no se pueden visualizar sin desnaturalizarlos.

---

[17] Con la posible e importante excepción de los países comunistas liderados por la antigua Unión Soviética, en los que una Filosofía del Estado (Diamat) basada en un desarrollo del materialismo dialéctico decimonónico evitó ciertos cuestionamientos y dio respuesta a los demás.

[18] L. von Bertalanffy, "*Das Gefüge des Lebens*", Teubner, Leipzig, 1937.

[19] L. von Bertalanffy, "*General system theory*", Penguin Books, Londres, 1968, pp. 255-261.

[20] No solo las cualidades secundarias van desapareciendo del contenido de la ciencia, sino también las cualidades primarias de la física clásica.

Por un lado, con esto quedan definitivamente fuera de la ciencia parcelas enteras de la realidad. El mundo del artista y el místico, así como el mundo de las vivencias de la persona común, parece alejarse de forma definitiva.
Pero por otro lado, como al final del camino la física formalizada no contendría nada específicamente humano, cabría esperar con Bertalanffy que el contenido matemático que expresa las relaciones físicas se pueda trasladar, mediante reglas de correspondencia adecuadas, resultando comprensible para seres inteligentes de otro planeta cuya experiencia sensorial se basara, por ejemplo, en la percepción de rayos X y hubieran alcanzado un nivel de abstracción comparable.
En este contexto adquiere pleno sentido el siguiente comentario que Max Delbrück realiza desde la perspectiva que brinda una epistemología evolutiva, "…la lección principal de las teorías general y especial de la relatividad es la siguiente: los seres humanos son organismos que pueden manipular sus representaciones internas del mundo mediante operaciones concretas y traspasar los límites de la percepción que les ha sido dada biológicamente. Pueden liberarse a sí mismos y construir una visión de la realidad que choca con la intuición, pero proporciona una visión más verdadera y de mayor alcance" (*Mente y Materia*, p. 297, Alianza Editorial, Madrid, 1986).

2. La mecánica newtoniana de partículas y su principio de relatividad

"No defino el tiempo, el espacio, el lugar y el movimiento, pues son bien conocidos de todos. Sólo debo observar que el vulgo concibe esas magnitudes como no sujetas a otras nociones, sino por la relación que guardan con objetos sensibles. Y de aquí surgen ciertos prejuicios para cuya eliminación será conveniente distinguirlos en absolutos y relativos, verdaderos y aparentes, matemáticos y comunes. El tiempo absoluto, verdadero y matemático, en sí mismo y por su propia naturaleza, fluye uniformemente sin consideración a nada externo, y también se llama duración; el tiempo relativo, aparente y común es alguna medida sensible externa (exacta o irregular) de la duración por medio del movimiento, y comúnmente se lo usa en vez del tiempo verdadero; por ejemplo una hora, un mes, un año."
Isaac Newton, Philosophiae Naturalis Principia Mathematica, Escolio a las definiciones iniciales, citado por Betrand Russell en El conocimiento humano, Parte IV, capítulo V, pp. 276-277, Planeta-Agostini, Barcelona, 1992.

"Para Newton, el espacio, como el tiempo, era "absoluto", es decir, consistía en una colección de puntos, cada uno desprovisto de estructura y siendo uno de los constituyentes últimos del mundo físico. Cada punto era eterno e inmutable; el cambio consistía en estar "ocupado" a veces por un trozo de materia, a veces por otro, y a veces por nada."
Betrand Russell en El conocimiento humano, Parte IV, capítulo VI, p. 287, Planeta-Agostini, Barcelona, 1992.

"Newton pensaba que el espacio no es solamente una más de las entidades físicas tradicionales; creía que el espacio es divino, convirtiendo así la geometría en una ciencia teológica. En consecuencia, los fundamentos axiomáticos de la geometría deberían revelar las leyes de la creación divina. En vista de esta perspectiva, uno puede hacerse una idea del ideal newtoniano de rigor matemático. En los últimos años de su vida, en la segunda edición de su principal obra, los Philosophiae Naturalis Principia Mathematica, Newton reiteró su reverencia hacia el espacio. Escribió: "Dios dura eternamente y está presente en todas partes; y al existir siempre y en todas partes, constituye la duración y el espacio…En El están y se mueven todas las cosas."…..
Las teorías de Newton unificaron la mecánica terrestre y la mecánica celeste por primera vez en la historia. De hecho, el concepto de fuerza que actúan a distancia es, a primera vista, una idea pintoresca que parece sospechosamente cercana a la magia. Pero Newton pertenecía a una época creyente y pensaba que el objetivo de la ciencia es mostrar cómo Dios ha hecho el mundo. El Dios de Newton era el benevolente creador bíblico, que puso en marcha el reloj celestial al comienzo de los tiempos y presumiblemente le da un vistazo de vez en cuando, para comprobar que la acumulación de perturbaciones gravitatorias no se desmanda."
Max Delbrück, Mente y materia, capítulo 14, Alianza Editorial, Madrid, 1986.

La mecánica newtoniana o mecánica clásica es un marco de conceptos y leyes adecuadas para describir el movimiento de los cuerpos materiales tales como las máquinas, los automóviles, las partículas de polvos y los coloides: es una mecánica de las masas intermedias. Con algunas salvedades (casos que requieren aplicar la teoría de

la gravitación basada en la relatividad generalizada), resulta extensible a la descripción del movimiento de los planetas del sistema solar, los cometas y el polvo cósmico. Algunas veces la mecánica newtoniana se puede aplicar a los movimientos de traslación de moléculas e incluso de átomos, pero es inaplicable al movimiento de los electrones.
Como se dijo en la introducción, cuando las dimensiones de un cuerpo material se pueden ignorar considerándolo como localizado en cada instante en un punto del espacio, se tiene una partícula. En mecánica clásica se asume que un cuerpo material extenso se puede reducir a un conjunto de partículas componentes. Si durante el movimiento las distancias entre esas partículas componentes se mantienen invariantes, el cuerpo se denomina rígido[21]. De lo contrario, se trata de un cuerpo deformable cuyo estudio debe efectuarse recurriendo a una extensión no trivial de la mecánica de partículas y cuerpos rígidos, la mecánica de los medios continuos.

Comencemos por revisar brevemente algunos aspectos de interés para comprender la elaboración de conceptos conducente a la mecánica de Newton.
Hasta el renacimiento se admitía la venerable doctrina aristotélica según la cual la velocidad de un cuerpo es proporcional a la fuerza aplicada, con su misma dirección y sentido. Cuando los cuerpos ascendían, como el humo, se suponía que actuaba una fuerza de levedad, y cuando caían al suelo, como una piedra, una fuerza de gravedad.
El abandono del sistema geocéntrico por el heliocéntrico hizo posible pensar de una forma que hasta ese momento no había sido posible. Por ejemplo, si admitimos que la Tierra gira alrededor del Sol, y la fuerza responsable del movimiento tuviera que ver con su velocidad, entonces sería una fuerza tangencial a la órbita terrestre. Si la fuente de esa fuerza estuviera en los cuerpos celestes, podríamos mirar en la dirección de avance de nuestro planeta para ver qué encontramos. No encontramos nada en particular: a veces una estrella lejana, a veces otra, a veces nada que podamos detectar. Pero si miramos en una dirección aproximadamente perpendicular a la velocidad de la Tierra en su órbita, en la dirección y sentido de su aceleración, hallamos en todo

---

[21] Los cuerpos sólidos son siempre en alguna medida deformables, y las perturbaciones mecánicas se propagan en forma de ondas elásticas. Las velocidades de propagación de las ondas elásticas aumentan a medida que aumenta el módulo de rigidez del sólido, es decir a medida que disminuye la deformación del cuerpo frente a una misma solicitación mecánica. Si suponemos que en principio no existe cota superior para estas velocidades de propagación, podemos concebir una secuencia de sólidos con módulos de rigidez crecientes tales que en el límite se obtendría, como idealización, un cuerpo rígido.
De acuerdo con la teoría de la relatividad, una modificación producida en un lugar del espacio físico solamente puede afectar lo que sucede en un lugar remoto luego de un tiempo de propagación cuyo valor mínimo está acotado por el cociente entre la distancia que los separa dividida por una velocidad límite, la misma respecto de todos los marcos de referencia admisibles e igual a la velocidad de la luz en el vacío.
La existencia de una cota superior a la velocidad de propagación de las perturbaciones implica que no es posible imaginar una secuencia de sólidos cuyo módulo de rigidez aumente sin límite de modo que las velocidades de las ondas elásticas tienden a infinito. Entonces, a diferencia de lo que se supone en mecánica newtoniana, en principio los cuerpos materiales no se pueden idealizar mediante el concepto de cuerpo rígido, puesto que una acción exterior que modifique la posición de uno de sus puntos no puede modificar simultáneamente la posición de los otros puntos de modo de conservar las distancias entre ellos. Así pues, si se admite una cota superior para la velocidad de propagación de las perturbaciones en los cuerpos materiales, hay dos tipos de cuerpos: los cuerpos deformables y las partículas, que por definición no poseen extensión. Un cuerpo deformable cuyas dimensiones son muy pequeñas respecto de las escalas espaciales en las que varían significativamente las fuerzas y se produce el movimiento, puede ser considerado como partícula
No obstante, la idealización del comportamiento mecánico de los sólidos mediante cuerpos rígidos conserva toda la importancia, la utilidad y las limitaciones que siempre ha tenido en la práctica.

momento un mismo cuerpo dominante: el Sol. Si suponemos que de alguna manera el Sol produce una fuerza que causa el movimiento de la Tierra en su órbita, entonces podría conjeturarse que la fuerza cambia la dirección de la velocidad: produce aceleración, no velocidad.
Generalizando a partir de este caso particular, en vez de una proporcionalidad aristotélica entre fuerza y velocidad, Newton conjeturó la proporcionalidad entre fuerza y aceleración, tanto para los movimientos celestes como para los terrestres.
Como la aceleración se determina a partir del cambio en la velocidad y el cambio en la velocidad a partir del cambio en la posición del cuerpo, es necesario determinar la posición a medida que transcurre el tiempo. La determinación de la posición de un cuerpo se puede realizar tomando otro cuerpo como referencia.
La relación entre fuerza y aceleración implicaba la ley de inercia, concebida ya por Galileo y otros antes que él: un cuerpo no sometido a fuerzas debidas a la interacción con otros cuerpos se mantendría en reposo o se movería con movimiento rectilíneo uniforme, tendría aceleración nula.
Sí, pero ¿aceleración nula respecto de qué? Porque dependiendo de qué cuerpo se elija para determinar las posiciones, la aceleración de un mismo cuerpo puede tomar un valor cualquiera. Por ejemplo, supongamos que un pasajero se desplaza en un automóvil en un tramo recto de carretera con velocidad constante respecto del suelo. En esas condiciones posee aceleración nula tanto respecto del vehículo como respecto de la tierra, pero se encuentra acelerado respecto del sol, y como el sol se encuentra acelerado respecto del centro de nuestra galaxia… Si hay una relación objetiva entre aceleración y fuerza de interacción entre los cuerpos, entonces se verifica solo si se elije adecuadamente el marco de referencia. Pero, como veremos a continuación, esta elección implica una elección de un proceso de medida del tiempo y de un proceso de medida de longitudes.

2.1 Marcos de referencia inerciales: reglas y relojes newtonianos

Para una descripción cuantitativa del movimiento, en mecánica newtoniana de partículas y cuerpos rígidos se utiliza una clase de marcos de referencia denominados **marcos inerciales**, que permiten ubicar la posición (relativa al marco de referencia) de cada partícula en todo instante de tiempo, dando para cada instante sus tres coordenadas espaciales.
En la práctica los marcos de referencia se construyen a partir de cuerpos que se consideran rígidos, utilizando reglas para medir distancias a partir de un origen solidario con el cuerpo, y relojes para medir el tiempo.
Para medir el tiempo en física se utiliza una clase de relojes particular, de los cuales un reloj de péndulo es el ejemplo más simple. Se denominan "**relojes newtonianos**", para diferenciarlos de los "no newtonianos", como un reloj de agua. Éste mide el intervalo de tiempo entre dos sucesos por el volumen de líquido que abandona un recipiente a través de un orificio en su parte inferior desde que ocurre el primer suceso y hasta que ocurre el segundo suceso. A medida que queda menos agua en el recipiente, sale más lentamente respecto del tiempo determinado por un reloj de péndulo: el volumen correspondiente a un mismo número de oscilaciones del péndulo es cada vez menor. Empleando un reloj de péndulo como base de comparación, el reloj de agua atrasa cada vez más y la duración que el reloj de agua asigna al intervalo de tiempo entre dos eventos separados por el mismo número de oscilaciones del reloj de péndulo es cada vez menor. En el capítulo sobre la teoría generalizada de la relatividad volveremos sobre este asunto.

En mecánica clásica el tiempo es absoluto e independiente del espacio. El instante en el que se produce un evento (por ejemplo el choque de dos partículas o la desintegración de un núcleo radioactivo) es el mismo en todos los marcos de referencia. Se representa mediante un número real. Un instante es posterior a otro cuando el número real que le corresponde a este último es mayor que el número real que corresponde al primer instante. Los relojes permiten establecer correspondencias entre instantes y números. Se supone que los relojes se pueden sincronizar, independientemente de su ubicación y de su estado de movimiento relativo a un marco de referencia, de modo tal que asignen al mismo instante de tiempo absoluto el mismo número real.

Las coordenadas espaciales del evento son relativas, es decir dependen del marco de referencia utilizado en la descripción del movimiento. Por tanto la cinemática (la descripción del movimiento utilizando conceptos tales como trayectorias, velocidades y aceleraciones) de un sistema de partículas se debe efectuar en cada caso respecto de un marco de referencia dado.

Si se emplea el concepto de espacio absoluto de Newton, uno de los marcos inerciales se encuentra fijo en el espacio absoluto, mientras que los demás marcos inerciales se encuentran en movimiento rectilíneo y uniforme respecto del espacio absoluto.
Se asume que la geometría de Euclides, revisada desde la perspectiva del siglo XVII, constituye el modelo matemático que permite describir este espacio absoluto, considerado como un continente amorfo y auto-conservado de dimensiones infinitas.
Se admitía que los cuerpos materiales se podían poner y sacar sin alterar la naturaleza del continente. En palabras de Newton[22]: “Podemos prescindir mentalmente de las cosas, pero no del espacio que ocupan”.

2.2 Las leyes de la mecánica

Respecto de los marcos de referencia inerciales las leyes de la mecánica adoptan la misma forma y la más simple.
Para plantear estas leyes se puede comenzar con el concepto de masa. En este dominio de las masas medias, se podría definir en forma tentativa la masa de un cuerpo a través de su peso a nivel del suelo.
No obstante, la masa que aparece en la formulación de las leyes de la mecánica clásica es la denominada **masa inercial**, mientras que la masa que aparece en relación con el peso de un cuerpo es la denominada **masa gravitatoria**. Los experimentos más precisos sugieren que la masa inercial y la masa gravitatoria de un cuerpo pueden considerarse proporcionales (y se pueden definir de modo que sean numéricamente iguales). Pero en principio se trata de conceptos diferentes, como lo comprendió y formuló el capitán de arqueros y senador genovés Juan Bautista Balliani, cuando en 1646 introdujo la distinción entre **moles** y **pondus**.
Newton mismo definió la masa inercial como “la cantidad de materia de un cuerpo medida por el producto de su densidad y su volumen”, lo cual, al igual que su definición de fuerza como “cualquier acción sobre un cuerpo que cambia o tiende a cambiar su estado de reposo o de movimiento en línea recta”, no dejan de presentar cierta circularidad desde el punto de vista lógico, como lo comprendió y explicó claramente Ernst Mach en 1883 en la primera edición de su influyente tratado de mecánica clásica.

[22] J. Singh “*Teorías de la cosmología moderna*”, p. 143, Alianza, Madrid, 1974.

Mach eliminó esa circularidad lógica analizando los resultados de experimentos realizados con pares de cuerpos que actúan recíprocamente el uno sobre el otro, como cuando se encuentran unidos por un resorte o sometidos a atracción gravitatoria mutua. Los resultados de los experimentos sugieren que el cociente de las magnitudes de las aceleraciones opuestas que de esta interacción mutua resultan, es constante y depende solamente de algo que hay en los cuerpos y que se puede denominar **inercia** o **masa inercial**. El cociente de las aceleraciones se puede utilizar para definir el cociente de las masas, en general diferentes, m_1 y m_2 : $\frac{m_2}{m_1} = \frac{a_1}{a_2}$

Eligiendo una masa como unidad, a partir de este cociente se pueden definir las masas correspondientes a distintos cuerpos materiales sin recurrir a la definición gravitacional, basada en el peso.

La fuerza de interacción entre cuerpos se puede definir a su vez mediante el producto de la masa por la aceleración de cada cuerpo.

Puesto que el concepto de espacio absoluto de Newton tiene mucho de hipótesis metafísica, según se desprende de las citas que dan inicio a la presente sección, Mach propuso sustituirlo por el sistema formado por las estrellas fijas y definir los marcos inerciales como aquellos en movimiento rectilíneo uniforme respecto del marco correspondiente a las estrellas fijas, cuando se emplean relojes newtonianos para medir el tiempo.

Ahora bien, si dotamos de cierta velocidad inicial a un cuerpo, que se encuentra en un ambiente bien definido y con una posición inicial conocida, se desprende del marco conceptual de la mecánica clásica que su movimiento futuro está determinado: se desplaza, rota y se deforma de manera enteramente previsible. Con ambiente bien definido, con influencia sobre la trayectoria, se hace referencia a la totalidad de las fuerzas que actúan sobre el cuerpo.

2.2.1 Primera, segunda y tercera ley de Newton

Si el cuerpo se mueve en el vacío sometido a la atracción gravitatoria de la Tierra, y no está sometido a otras fuerzas, su trayectoria resulta independiente de su masa. Supongamos ahora que el cuerpo posee carga eléctrica, se mueve exclusivamente debido a fuerzas electro-estáticas y la descripción de su movimiento se lleva a cabo respecto de un marco de referencia inercial.

Si la carga y las dimensiones del cuerpo se mantienen constantes pero la masa aumenta, la desviación de la partícula producida por las fuerzas se hace cada vez menor hasta que en el límite de masa infinita el cuerpo se mueve en forma rectilínea y uniforme respecto de un sistema inercial. Esta es la **primera ley** de la mecánica clásica.

En general, para una partícula (cuya posición en el instante t viene dada por un vector $\vec{r}(t)$ relativo a un marco inercial), el producto de su masa m por su aceleración $\vec{a}(t) = \frac{d\vec{v}(t)}{dt} = \frac{d^2\vec{r}(t)}{dt^2}$ en cada punto del espacio se mantiene constante si el ambiente no varía, o sea si la suma vectorial de las fuerzas $\sum_j \vec{F}_j$ que actúan sobre la partícula no varía.

La **segunda ley** de la mecánica clásica establece que la aceleración, la masa inercial y la fuerza resultante verifican: $m \cdot \vec{a} = \sum_j \vec{F}_j$

La **tercera ley** de la mecánica clásica establece que cuando dos cuerpos interactúan, la fuerza que uno de ellos ejerce sobre el otro cuerpo es igual y opuesta a la fuerza que este último ejerce sobre el primero.

Si las fuerzas que actúan sobre la partícula se conocen en todo instante y en todo punto del espacio y si se fijan la posición en un instante tomado como inicial (un vector de posición) y la velocidad inicial para la partícula, la segunda ley de Newton permite tanto pre-decir como retro-decir la trayectoria.

2.2.2 Conservación de la cantidad de movimiento y del momento angular de un sistema de partículas. Centro de masas.

El análisis del movimiento de las partículas y de los sistemas de partículas se facilita considerablemente cuando se pueden aplicar principios de conservación de la cantidad de movimiento y de la energía.

La cantidad de movimiento de una partícula es un vector $\vec{p}$ que se define como el producto de su masa (escalar) y su velocidad (vector). Para un sistema se define como la suma vectorial de las cantidades de movimiento de las partículas que lo componen: $\vec{p}_s = \sum_j \vec{p}_j$ La variación en la cantidad de movimiento de un sistema se relaciona con la resultante de las fuerzas que actúan sobre las partículas componentes: $\frac{d}{dt}\vec{p}_s = \sum_j \vec{F}_j$

La suma de las fuerzas de interacción entre las partículas del sistema, debido a la tercera ley, es nula. Entonces la variación en $\vec{p}_s$ se debe a la resultante de las fuerzas externas.

Si un sistema se encuentra aislado, es decir, no sometido a otras fuerzas que las de interacción entre las partículas que lo componen, entonces la cantidad de movimiento del sistema permanece constante.

Si $\vec{r}_1, \vec{r}_2, ..., \vec{r}_N$ son los vectores de posición en un instante dado de un sistema de N partículas de masas $m_1, m_2, ..., m_N$ respectivamente, entonces se define el centro de masas del sistema como un punto con vector de posición $\vec{r}_c = \frac{1}{\sum_{j=1}^{N} m_j} \sum_{j=1}^{N} m_j \cdot \vec{r}_j$

De su definición se desprende que el centro de masas se mueve con una velocidad $\vec{v}_c = \frac{1}{\sum_{j=1}^{N} m_j} \sum_{j=1}^{N} \vec{p}_j$ Si la cantidad de movimiento del sistema permanece constante, entonces su centro de masa se mueve con velocidad constante.

Suponiendo un choque de dos partículas de masas m_1 y m_2 que en general produce otras dos partículas de masas m_3 y m_4 (como sería el caso de una reacción química), la conservación de la cantidad de movimiento se formula así, en términos de las cantidades de movimiento inmediatamente antes e inmediatamente después del choque:

$$\vec{p}_1 + \vec{p}_2 = m_1 \cdot \vec{v}_1 + m_2 \cdot \vec{v}_2 = m_3 \cdot \vec{v}_3 + m_4 \cdot \vec{v}_4 = \vec{p}_3 + \vec{p}_4$$

En mecánica clásica las masas se conservan, de modo que: $m_1 + m_2 = m_3 + m_4$

Eligiendo un punto como origen, el **momento angular** o **momento de la cantidad de movimiento** de una partícula respecto de ese punto se define como el producto vectorial del vector de posición con la cantidad de movimiento en un mismo instante: $\vec{L} = \vec{r} \times m \cdot \vec{v}$

El momento de una fuerza se define como el producto vectorial $\vec{M} = \vec{r} \times \vec{F}$ De estas definiciones y de la segunda ley de Newton $m \cdot \frac{d}{dt}\vec{v} = \vec{F}$ se desprende una relación entre el momento de la fuerza y la variación del momento angular: $\frac{d}{dt}\vec{L} = \vec{M}$

El momento angular de un sistema de partículas se define como la suma vectorial de los momentos angulares de las partículas componentes, tomados en el mismo instante respecto del mismo punto: $\vec{L}_s = \sum_j \vec{L}_j$

Sumando los momentos de las fuerzas que actúan sobre las partículas del sistema se obtiene un momento de fuerza resultante $\vec{M}_s = \sum_j \vec{M}_j$ Las contribuciones debidas a las fuerzas de interacción entre las partículas que componen el sistema se anulan unas a otras, como consecuencia de la tercera ley, de modo que el momento se reduce a la suma de momentos debidos a las fuerzas externas al sistema.

Para el sistema considerado como un todo se verifica: $\frac{d}{dt}\vec{L}_s = \vec{M}_s$

La **energía cinética** de una partícula se define como la mitad del producto de su masa y el cuadrado de la magnitud de su velocidad.

La energía cinética de un sistema de partículas se define como la suma de las energías cinéticas de esas partículas.

Cuando la fuerzas de interacción entre la partícula y su ambiente, o entre las partículas de un sistema entre sí y con su ambiente, poseen ciertas propiedades que permiten asimilarlas a los denominados campos de fuerza conservativos, se puede introducir la denominada energía potencial y formular un **teorema de conservación de la energía**: la energía y su conservación se consideran en el capítulo 3, destinado a analizar el concepto de campo.

Para el caso de un choque de dos partículas la conservación de la energía se puede formular a partir de las energías cinéticas inmediatamente antes e inmediatamente después del choque. La energía Q se introduce para describir choques inelásticos:

$$\frac{1}{2} \cdot m_1 \cdot v_1^2 + \frac{1}{2} \cdot m_2 \cdot v_2^2 = \frac{1}{2} \cdot m_3 \cdot v_3^2 + \frac{1}{2} \cdot m_4 \cdot v_4^2 + Q$$

2.3 El principio de relatividad de la mecánica clásica

Un **principio de relatividad** se encuentra asociado con la mecánica clásica: el principio de relatividad que se suele denominar principio de Galileo. Este principio conecta la descripción cinemática de los sucesos respecto de un marco inercial K con la descripción cinemática de los sucesos respecto de otro marco inercial cualquiera K' de tal manera que la formulación matemática de las tres leyes de la mecánica clásica permanece invariantes al pasar de un marco al otro.

En el caso más sencillo, se emplean sistemas de coordenadas cartesianas en ambos marcos de referencia, y el movimiento de K' respecto de K es una traslación a lo largo del eje de las x con velocidad V en condiciones tales que los orígenes de los sistemas de coordenadas coinciden en el instante inicial, el mismo en ambos marcos. En ese caso un evento (por ejemplo, la desintegración de un núcleo radiactivo) que en el instante t se produce en el punto de coordenadas x, y, z respecto de K, respecto de K' se produce en un instante t' y presenta coordenadas x', y', z' dadas por las denominadas **transformaciones de Galileo**:

$$t' = t \qquad x' = x - V \cdot t \qquad y' = y \qquad z' = z$$

La trayectoria de una partícula respecto de K viene dada por la ley horaria $x(t)$ $y(t)$ $z(t)$. Entonces las velocidades, respecto de K, del movimiento de una partícula $\frac{dx(t)}{dt}$ $\frac{dy(t)}{dt}$ $\frac{dz(t)}{dt}$, se vinculan con las velocidades del movimiento respecto de K' a través de las ecuaciones siguientes, que expresan la ley de adición de las velocidades clásica:

$$v_x' = \frac{dx'(t')}{dt'} = \frac{dx(t)}{dt} - V = v_x - V \qquad v_y' = \frac{dy'(t')}{dt'} = \frac{dy(t)}{dt} = v_y \qquad v_z' = \frac{dz'(t')}{dt'} = \frac{dz(t)}{dt} = v_z$$

Con mayor generalidad, si $\vec{r}$ es el vector de posición de un evento respecto de K, si $\vec{r}'$ es su vector de posición respecto de K' y si $\vec{V}$ es la velocidad de K' respecto de K, entonces las transformaciones de Galileo adoptan la expresión vectorial:

$$t' = t \qquad \vec{r}' = \vec{r} - t \cdot \vec{V}$$

De las transformaciones de Galileo se desprende que la diferencia entre los vectores de posición de dos partículas en un mismo instante es la misma respecto de todos los marcos de referencia: $\Delta\vec{r}' = \Delta\vec{r}$

Como consecuencia la distancia entre ellas es la misma. Entonces, en la mecánica newtoniana la longitud de una varilla es independiente del marco de referencia respecto del cual se la mide: es absoluta

.

La trayectoria de una partícula respecto de K viene dada por la ley horaria $\vec{r}(t)$ mientras que respecto de K' viene dada por la ley horaria $\vec{r}'(t') = \vec{r}'(t)$.

A partir de las transformaciones de Galileo se desprende que las velocidades de esa misma partícula $\vec{v}(t) = \frac{d}{dt}\vec{r}(t)$ y $\vec{v}'(t) = \frac{d}{dt}\vec{r}'(t)$ se relacionan a través de la **ley de adición de las velocidades de la mecánica clásica**:

$$\vec{v}'(t) = \vec{v}(t) - \vec{V} \qquad \vec{v}(t) = \vec{v}'(t) + \vec{V}$$

En la medida en que todos los marcos inerciales se consideren equivalentes, el concepto de reposo absoluto carece de sentido. Si un cuerpo se encuentra en reposo respecto de K, estará en movimiento respecto de K', y no hay razón para preferir uno de ellos. La noción de velocidad absoluta también carece de sentido: solo la velocidad relativa de un cuerpo respecto de otro posee significado físico.

No obstante, en mecánica newtoniana tiene sentido una aceleración absoluta.

Como $\vec{V}$ es constante, las aceleraciones $\vec{a}(t) = \frac{d}{dt}\vec{v}(t)$ respecto de K y $\vec{a}'(t) = \frac{d}{dt}\vec{v}'(t)$ respecto de K' son iguales. Entonces **la aceleración de una partícula es la misma**

respecto de cualquier marco de referencia inercial. Dicho de otra forma, es un invariante bajo las transformaciones de Galileo.
En mecánica clásica la masa inercial se asume independiente del estado de movimiento del cuerpo, así que también es un invariante.
Como las fuerzas de interacción a distancia entre las partículas materiales se supone son las mismas en cualquier marco de referencia, la ecuación que expresa la segunda ley $m \cdot \vec{a} = \sum_j \vec{F}_j$ es invariante bajo las transformaciones de Galileo.

Consideremos ahora un marco de referencia no inercial Ñ que se mueve con velocidad $\vec{V}(t)$ con respecto de un marco inercial K. En ese caso la velocidad $\vec{v}(t)$ de un cuerpo respecto de K se puede expresar como la suma de su velocidad $\vec{v}'(t)$ respecto del marco no inercial y la velocidad de este último respecto de K: $\vec{v}(t) = \vec{v}'(t) + \vec{V}(t)$
Si $m \cdot \vec{a} = \vec{F}$ es la expresión de la segunda ley de Newton respecto de K, entonces respecto de Ñ la segunda ley adopta la forma:

$$m \cdot \vec{a}' = \vec{F} - m \cdot \vec{W}$$

Aquí $\vec{a}' = \dfrac{dv'(t)}{dt}$ es la aceleración del cuerpo relativa al marco no inercial Ñ y $\vec{W}(t) = \dfrac{d\vec{V}(t)}{dt}$ es la aceleración de Ñ relativa a K. Respecto de un sistema no inercial el movimiento se produce bajo una fuerza adicional, proporcional a la masa del cuerpo cuyo movimiento se considera, posiblemente variable en el tiempo y en el espacio. Se la denomina **fuerza de inercia**.
Si Ñ se traslada y rota respecto de K, la aceleración $\vec{W}(t)$ se puede descomponer en la suma de cuatro términos: una aceleración de arrastre asociada a una traslación no uniforme de Ñ respecto de K, y tres términos debidos a la rotación de Ñ respecto de K. Dos de ellos involucran la velocidad angular del marco no inercial respecto del marco inercial: una aceleración centrífuga y una aceleración de Coriolis. El tercero involucra la aceleración angular.
La aparición de las fuerzas de inercia $-m \cdot \vec{W}$ establece una diferencia entre los marcos inerciales y los no inerciales que es absoluta.

Al respecto dice Newton en el Escolio[23] a los "Principios matemáticos de la filosofía natural":
"Los efectos que distinguen el movimiento absoluto del relativo son las fuerzas que actúan alejándose del eje en un movimiento circular. Pues tales fuerzas no existen en un movimiento circular puramente relativo, mientras que en un movimiento circular verdadero y absoluto son mayores o menores según la cantidad de movimiento.
Si de una cuerda de buena longitud suspendemos un cubo, lo hacemos girar hasta que la cuerda se halle fuertemente retorcida y lo llenamos luego de agua, manteniéndolo en reposo junto con el agua, entonces, por una acción repentina de otra fuerza, el cubo comenzará a girar en sentido contrario, y mientras la cuerda recupera su estado primitivo, el cubo persistirá durante algún tiempo en este movimiento; al principio la superficie del agua será plana, tal y como lo era antes de que el cubo comenzara a girar,

[23] Ver Einstein y otros "*La teoría de la relatividad*", Altaya, Barcelona, 1993, pp.21-22.

más después el cubo, al ir comunicando gradualmente su movimiento al agua, obligará a ésta a iniciar un movimiento sensible de rotación y a alejarse poco a poco del centro y ascender por las paredes del recipiente, tomando la forma de una figura cóncava (como yo mismo he experimentado), y cuanto más rápido se hace el movimiento, tanto más subirá el agua, hasta que, por último, ejecutando sus revoluciones al unísono con el recipiente, quedará en reposo relativo dentro de él. Esta ascensión del agua demuestra su propensión a alejarse del eje del movimiento; y el movimiento circular verdadero y absoluto del agua, que aquí es directamente contrario al relativo, se pone de manifiesto y puede ser medido a través de dicha propensión."
Siguiendo esta misma línea de razonamiento se puede explicar la forma achatada en los polos de nuestro planeta: si éste gira respecto del espacio absoluto alrededor de su eje, entonces aparecerán fuerzas centrífugas que tienden a dilatarlo tanto más cuanto más próximo se encuentre la zona considerada del ecuador y al mismo tiempo la aceleración de la gravedad será menor donde la fuerza centrífuga sea mayor, cosa que efectivamente ocurre.

En "La ciencia de la mecánica" Mach[24] retoma el argumento del cubo de agua de Newton, y dice al respecto:
"…si fundamentamos nuestros argumentos sobre la base de los hechos comprobamos que el único conocimiento posible es el de espacios y movimientos **relativos**. **Relativamente**, o lo que es lo mismo, haciendo caso omiso de ese ignoto e ignorable medio del espacio, los movimientos del universo son los mismos independientemente de que adoptemos la concepción ptolemaica o la copernicana. Ambas concepciones son, de hecho, igual de **correctas**, sólo que la última es más simple y más **práctica**. El universo no nos ha sido dado **dos veces**, una con una tierra en reposo y otra vez con una tierra en movimiento, sino **una sola vez**, con sus movimientos **relativos**, que es lo único susceptible de ser determinado. …El experimento de Newton con el cubo de agua en rotación, únicamente nos dice que la rotación relativa del agua respecto a las paredes del recipiente no produce **ninguna** fuerza centrífuga perceptible, pero que tales fuerzas **sí se producen** como consecuencia de la rotación relativa respecto de la masa de la tierra y demás cuerpos celestes. Y nadie está autorizado para predecir qué ocurriría si las paredes del recipiente aumentaran de espesor y de masa hasta alcanzar un espesor de varias leguas…Las consideraciones que acabamos de exponer demuestran que no es necesario referir la ley de inercia a ningún espacio absoluto especial. Todo lo contrario: las masas que, según la fraseología común, ejercen fuerzas unas sobre otras y aquellas que no ejercen para nada dichas fuerzas, guardan una relación muy similar respecto de la aceleración. De hecho, podemos considerar que **todas** las masas están relacionadas entre sí."
En lugar del espacio absoluto, Mach propuso utilizar como **marco inercial básico** el sistema que ya era empleado en astronomía, fijo al sistema estelar considerado como un todo. Los demás marcos inerciales se definen como en movimiento rectilíneo y uniforme en relación con este sistema fijo a las estrellas.
Desde el punto de vista de Mach la dinámica del universo debería describirse mediante las modificaciones en las distancias relativas entre los cuerpos, el tiempo debería deducirse de ese cambio y el movimiento inercial podría ser la consecuencia de la acción de la totalidad de las masas del Universo sobre el cuerpo considerado.

---

[24] Ver Einstein y otros "*La teoría de la relatividad*", Altaya, Barcelona, 1993, pp.30-32.

Este programa fue el principal estímulo para la construcción de la teoría generalizada de la relatividad por Einstein. El camino de Einstein se inicia con el análisis de las transformaciones que dejan invariantes las leyes de la naturaleza y el principio de relatividad se expresa a través de esa invariancia. Por este camino, el programa de Mach desemboca en una teoría del tiempo-espacio-materia, todavía inconclusa. Muchos creen que podría culminar, o que ya ha culminado, bajo la forma de una teoría del campo unificado.

Pero el programa de Mach puede abordarse siguiendo otro camino, planteado por J. H. Poincaré en 1902[25] e implementado a partir de la década de 1970 por J. Barbour [26] junto con un pequeño grupo de científicos.

Para ver cómo Poincaré planteó el problema de la relatividad del movimiento, consideremos un sistema aislado de N cuerpos que interactúan entre sí a través de fuerzas dirigidas a lo largo de los segmentos que unen los cuerpos dos a dos. Una buena aproximación a un sistema aislado de este tipo puede ser el sistema solar.

Si admitimos que la evolución del sistema se ve sometida a las leyes de la dinámica de Newton y a su ley de atracción gravitatoria, entonces alcanza con conocer las posiciones y las velocidades de cada cuerpo, en un instante inicial y respecto de un marco inercial, para que las trayectorias futuras puedan ser pre-dichas y las trayectorias pasadas retro-dichas en su totalidad.

No obstante, si nos limitamos a trabajar con la información que podemos obtener desde adentro del sistema, solo accederemos a las distancias mutuas entre los cuerpos y a la velocidad de variación de estas distancias.

Parecería que una mecánica compatible con el programa de Mach debería ser capaz de predecir el movimiento a partir de una condición inicial dada en términos de las distancias entre los cuerpos y sus velocidades de variación.

Pero la mecánica newtoniana exige más información, como lo muestra el siguiente ejemplo. Consideremos un sistema de dos cuerpos, por ejemplo el sol y un planeta, siendo r la distancia entre ambos y $\dot{r}$ su velocidad de variación. Cuando r pasa por un valor mínimo r_m, entonces $\dot{r} = 0$. Tomemos r_m, $\dot{r} = 0$ como estado inicial. Pero este estado inicial no alcanza para determinar la trayectoria del planeta respecto del Sol: es compatible con una órbita elíptica, con una hiperbólica y con una caída radial del planeta hacia el sol. Lo que falta es un dato que defina el momento angular de la órbita. Pero el momento angular no se puede hallar desde adentro del sistema, recurriendo a las distancias mutuas y a sus velocidades de variación solamente. Es necesario añadir información sobre las aceleraciones mutuas, o bien aumentar la información haciendo referencia a algo exterior al sistema, que en el caso de la mecánica newtoniana es el espacio absoluto u otro marco inercial.

La mecánica newtoniana no verifica el principio de relatividad en su versión fuerte debida a Poincaré, según el cual las distancias mutuas y sus velocidades de variación deberían alcanzar para determinar la evolución del sistema[27].

---

[25] J. H. Poincaré "*La ciencia y la hipótesis*", Espasa Calpe, Buenos Aires, 1943.

[26] J. Barbour "Relational concepts of space and time" The British Journal for the Philosophy of Science" vol.33, Nº 3, 1982, pp. 251-274; J. Barbour y B. Bertotti "Mach's principle and the structure of dynamical theories" Proceedings of the Royal Society London A, vol. 382, 1982, pp. 295-306; J. Barbour "The nature of time", 2009, arXiv:0903.3489 [gr-qc]; J. Barbour "The definition of Mach's principle", 2010, arXiv:1007.3368v1[gr-qc].

2.4 Atracción gravitatoria

Una relación general y particularmente importante entre masas, que en su momento motivó el desarrollo de la teoría generalizada de la relatividad, es precisamente la que establece la atracción gravitatoria entre ellas acompañada de la igualdad entre las masas inercial y gravitatoria de un mismo cuerpo.

Para la atracción gravitatoria entre dos masas gravitatorias puntuales $m_{g,1}$ y $m_{g,2}$, separadas por una distancia d_{12}, Newton planteó la siguiente expresión, en la cual G = 6,67 × 10^{-8} $cm^3/g \times s$ es la denominada constante de gravitación universal y el signo – indica su naturaleza siempre atractiva:

$$F_{1,2} = -G \cdot \frac{m_{g,1} \cdot m_{g,2}}{{d_{12}}^2}$$

Supongamos que a_1 es la aceleración que un campo gravitatorio g produce en el cuerpo de masa inercial m_1 (con masa gravitatoria $m_{g,1}$) y a_2 es la aceleración producida en el cuerpo con masa inercial m_2 (con masa gravitatoria $m_{g,2}$) por ese mismo campo. Por la segunda ley de Newton: $m_1 \cdot a_1 = m_{g,1} \cdot g \qquad m_2 \cdot a_2 = m_{g,2} \cdot g$

Entonces, siempre y cuando $a_1 = a_2$ para todo par de cuerpos de prueba sometido a un campo gravitatorio de la misma magnitud, se verificará para todo par de cuerpos la proporcionalidad entre la masa inercial y la masa gravitatoria $\frac{m_1}{m_{g,1}} = \frac{m_2}{m_{g,2}}$

En unidades adecuadas esa proporcionalidad se expresa como una igualdad.

La exactitud con la que en cada caso se puede comprobar la igualdad de las masas inercial y gravitatoria de un mismo cuerpo está determinada por la exactitud con que se pueda comprobar que $a_1 = a_2$.

La evidencia experimental[28] sugiere que la **masa gravitatoria y la masa inercial de un cuerpo son proporcionales, pudiendo tomarse como iguales**: $m_g = m$

Si bien el concepto de fuerzas que actúan a distancia parecía extraño (Newton mismo objetaba este concepto, por razones filosóficas), la aplicación de la mecánica de partículas y la ley de gravitación newtoniana al sistema solar y al movimiento de los cuerpos terrestres tuvo éxito asombroso, al igual que su extensión para describir la deformación de los sólidos elásticos y el flujo de líquidos y gases.

Además, el principio de relatividad de Galileo implica que las interacciones entre los cuerpos materiales deben propagarse instantáneamente **a través del espacio vacío**: la modificación en el estado de un cuerpo debe afectar de inmediato a los otros cuerpos que interactúan con el primero, no importa cuán alejados de él se encuentren dichos cuerpos.

---

[27] J. Barbour "The definition of Mach's principle", 2010, arXiv:1007.3368v1[gr-qc].

[28] Esta igualdad fue confirmada, a menos de un pequeño error experimental de una parte en 10^8, en una famosa experiencia efectuada por Ëotvös en 1890 utilizando un péndulo de torsión. Se ha retomado el tema varias veces desde entonces, con técnicas experimentales más sofisticadas y la confirmación se ha producido a menos de un error cada vez menor. Los experimentos Eöt-Wash confirman la igualdad a menos de una parte en 10^{12} y la misión espacial STEP (Satellite Testing of the Equivalent Principle), si tiene éxito, podría suministrar una confirmación a menos de una parte en 10^{18}.

Si así no fuera, y una perturbación en un cuerpo se propagara por el vacío con velocidad finita respecto de un marco K', entonces por la ley de adición de velocidades $\vec{v}(t) = \vec{v}'(t) + \vec{V}$ esa perturbación se propagaría con una velocidad diferente respecto de otro marco K. Pero entonces sería posible distinguir K de K' desde el punto de vista de la mecánica, lo cual a su vez contradice el principio de relatividad de Galileo.
Así pues, **la acción a distancia**, tal como aparece en la ley de gravitación de Newton y en la ley de acción y reacción, **es una consecuencia del principio de relatividad de la mecánica clásica**[29].

2.6 Mecánica de medios continuos y óptica ondulatoria

El concepto de partícula de la mecánica newtoniana desciende directamente de los átomos flotando en el vacío de Leucipo y Demócrito. A su vez la idea del átomo de los filósofos griegos parece haber surgido naturalmente al reflexionar sobre el modelo de materia de Anaxímenes.
A partir de observaciones de los objetos que lo rodeaban, este pensador formuló la conjetura siguiente: los estados de la materia son cuatro (sólido, líquido, gaseoso e ígneo) y cuando una porción de materia cambia de estado no altera su naturaleza sino que se expande o contrae según corresponda.
Al expandirse las partes de la materia se separan entre sí. Si suponemos que la materia se distribuye sin interrupción a través de un volumen, y luego se expande, ¿qué es lo que se separa?
Una forma de responder a ésta pregunta es añadir una nueva conjetura: la materia está formada por puntos aislados o parcelas diminutas invisibles, que se alejan unas de otras cuando una porción macroscópica (visible) de materia se expande, y se aproximan entre sí cuando se contrae. Si suponemos que estas rarefacciones y condensaciones no alteran la naturaleza de estos átomos, la suposición de Anaxímenes acerca de que la naturaleza de la materia no se altera al cambiar de estado se encuentra justificada.
No obstante esta fructífera, duradera y respetable tradición de atomismo, la aplicación de la teoría corpuscular de la materia basada en las ecuaciones de Newton para describir la deformación y el flujo de los materiales resultó inviable debido al gran número de corpúsculos involucrados. Esta descripción se necesitaba para hacer posible un ulterior desarrollo de disciplinas técnicas como la hidráulica y la resistencia de materiales. No se halló otro camino a seguir que promediar y efectuar aproximaciones. Se desarrollaron dos enfoques complementarios: el propio de la mecánica de los medios continuos, adaptado de inmediato al uso de los ingenieros, y el característico de la mecánica estadística, que permaneció durante varias décadas en manos de los físicos y en consecuencia tardó un poco más en integrarse al acervo de herramientas tecnológicas.
La mecánica de medios continuos supone que la materia se encuentra distribuida en los cuerpos tal como parece en una observación macroscópica: en forma continua. Se

[29] Esto no se contradice con la velocidad finita de propagación de las interacciones a través de un medio material continuo, como se observa en el caso de ondas mecánicas en sólidos u ondas acústicas en gases o líquidos, puesto que en estos casos la velocidad de propagación de las perturbaciones mecánicas se define en relación con un cuerpo material dado.

introduce un campo de desplazamientos para cada punto de un medio continuo. Se definen un estado local de deformaciones (deducido a partir del campo de desplazamientos del medio continuo) y un estado local de esfuerzos. Se adapta la formulación matemática de las leyes de la dinámica newtoniana de partículas a los medios continuos. A partir de los resultados de observaciones y experimentos cuidadosos se plantean relaciones constitutivas entre el estado de esfuerzos y el estado de deformaciones para los sólidos, y entre el estado de esfuerzos y la velocidad de cambio del estado de deformaciones para los líquidos y gases. Nacen así la teoría de la elasticidad, la hidrodinámica y la aerodinámica.
Al mismo tiempo se afirma el punto de vista ondulatorio acerca de la naturaleza de la luz. Fresnel, reflexionando acerca de las analogías entre la propagación de las ondas elásticas en los sólidos continuos y la propagación de la luz, conjeturó la existencia de un medio elástico que llenaría la totalidad del espacio y que sería el portador de las ondas luminosas. Las ondas luminosas se podrían reducir entonces a la propagación de deformaciones de ese medio, un éter elástico, según leyes mecánicas.

2.7 Mecánica clásica, causalidad eficiente, determinismo y libre albedrío

A partir de mediados del siglo XIX se extendió por Europa y su área de influencia intelectual, la idea de la materia como algo permanente, tangible en el espacio, dotado de masa, sujeto a leyes de evolución y conservación conocidas, leyes que podían ser y efectivamente eran corroboradas a partir de las observaciones y mediciones del movimiento.
Se hizo cada vez más común la creencia que la materia existía desde siempre, no pudiendo ser ni creada ni destruida.
En esta imagen del mundo físico la materia se movía de acuerdo con las leyes de Newton, cada porción impulsada por las fuerzas que sobre ella ejercían las demás porciones, e impulsándolas a su vez.
Para caracterizar el estado de un cuerpo en el sentido dinámico clásico del término, bastaba con especificar las posiciones y velocidades de sus partes constituyentes.
Muchas personas familiarizadas con las ciencias físico-matemáticas de aquel tiempo creían que si se conocían las posiciones y las velocidades iniciales de los diferentes cuerpos, es decir, sus estados dinámicos, entonces, por lo menos en principio, era posible predecir el movimiento en sus más mínimos detalles, recurriendo a fuerzas de interacción conocidas para dar cuenta de los cambios de estado de los cuerpos y de las transferencias de energía que acompañan a esos cambios. Esto a su vez permitía una interpretación de los movimientos en términos de cadenas causa-efecto.
A comienzos del siglo XIX, Laplace formula con bastante precisión este **determinismo científico** basado en el concepto de causa eficiente, considerado como doctrina sobre la estructura y dinámica del mundo.

En el siglo XIX y en el siglo XX el empleo del concepto de causa eficiente se extiende progresivamente de las ciencias físicas a todas las ciencias de la naturaleza, la psicología y la sociología. Esa extensión se acompaña de una reducción en el empleo de las causas finales, que terminan circunscritas apenas a algunas áreas de las ciencias humanas.

En el marco del paradigma de la mecánica clásica de partículas y medios continuos, se vuelve a plantear nuevamente en el siglo XIX el problema del libre albedrío, cuyos orígenes en la tradición de occidente se remontan a la filosofía griega, cinco siglos antes de Cristo.
Si la dinámica de la materia que forma el cuerpo de una persona está estrictamente determinada por leyes físicas y químicas que rigen sus interacciones con los cuerpos externos y entre las partes del organismo, los movimientos de sus extremidades también lo estarían. Siguiendo esta línea de pensamiento, parecería que los movimientos voluntarios de una persona no serían tales. Yendo un poco más lejos en la misma dirección, el sentimiento de ser capaz de elegir que todas las personas en mayor o menor grado en general presentan, y la correspondiente responsabilidad individual, podrían ser una ilusión.

Este problema preocupó a los primeros atomistas[30], en su intento de reducir el mundo y su dinámica a átomos y vacío.
Resurgió en un contexto teológico cuando San Agustín extrajo conclusiones sobre el libre albedrío razonando a partir de la omnisciencia y omnipotencia de Dios[31].
Los filósofos volvieron a plantear y a discutir el problema de la compatibilidad entre la libertad y el determinismo, una y otra vez, con ropajes diferentes según la época.

En el siglo XX se produce el apogeo y la posterior crisis del determinismo científico, cuestionado desde adentro de la física por el desarrollo de la mecánica cuántica en la primera mitad de ese siglo, y desde afuera de la física por el desarrollo de nuevos conceptos ontológicos y epistemológicos sobre sistemas y emergencia, en las ciencias y en la filosofía de la segunda mitad del siglo. No obstante el problema de la relación entre libertad y determinismo continuó ocupando un lugar en las discusiones filosóficas hasta la actualidad.

Aún en un marco de ideas decimonónicas, era posible distinguir los problemas de la libertad de los problemas del determinismo. Esta distinción elimina algunas dificultades del planteo como problema único y permite replantear toda la cuestión en relación con el libre albedrío y con el determinismo en otros términos.
Al parecer el primero que planteó esta distinción fue Carlos Vaz Ferreira en el Uruguay de comienzos del siglo XX[32].
Una diferencia notable entre la ciencia y la filosofía, sobre la que Vaz Ferreira insistía a menudo, se encuentra en el sentido que se le asigna a las palabras en una y otra de esas disciplinas.

---

[30] Se sabe que en particular preocupó a Demócrito, quien salió del brete sosteniendo que su modelo de átomos en el vacío era una conjetura destinada a explicar las percepciones de los sentidos, pero afirmaba, como un precursor de Kant, que ignoramos cómo son realmente las cosas (los noúmenos).

[31] De la omnisciencia y omnipotencia parecía desprenderse que las personas no pueden hacer algo que Dios no lo sepa y no solo que no lo consienta, sino que no lo determine.

[32] C. Vaz Ferreira, "*Estudios filosóficos. Antología*", Aguilar, Buenos Aires, 1961, pp. 21-58. Un enfoque en cierta medida análogo al de Vaz Ferreira, fue utilizado posteriormente por Ernst Cassirer (en "*Deteminism and indeterminism in modern physics*", Yale University Press, New Haven, USA, 1956) cuando discutió el problema del libre albedrío en relación con la crisis de la causalidad que acompañó la interpretación de Copenhague de la mecánica cuántica.

Cuando en los científicos decimonónicos se referían a la materia, esa palabra tenía más o menos el mismo significado para todos ellos. Partían de datos de la percepción, despojándolos de las propiedades que la psicología mostraba eran esencialmente de naturaleza subjetiva, pero pensaban sobre la materia más o menos en el mismo nivel de abstracción, utilizando el concepto en el marco de la mecánica clásica, dotándolo así de un significado implícito, pero sin discutir el concepto mismo. El nivel de abstracción era el mismo, por una especie de acuerdo tácito, lo cual permitía poner a prueba sin excesiva ambigüedad las proposiciones que involucraban el concepto físico de materia[33].

Los filósofos, por su parte, entre otras actividades, analizan los significados de los términos del lenguaje desde sus correspondientes posturas metafísicas y construyen sus teorías filosóficas.
A diferencia de lo que suelen hacer los científicos, los filósofos no trabajan siempre en un mismo nivel de abstracción, sino que se adentran a través de niveles de abstracción sucesivos.
A veces no explicitan el nivel de abstracción lo suficiente. Pero una vez comenzado el camino de abstracción progresiva, la misma palabra, por ejemplo materia, se utilizaba para aludir conceptos ubicados en niveles de abstracción diferentes, lo cual además de prestarse a una considerable confusión, puede tener como consecuencia que el valor de verdad de una afirmación en un nivel de abstracción puede modificarse en el nivel siguiente.
Esto en sí mismo no es un problema si se tiene presente el nivel de abstracción en el que se trabaja, pero lo es si éste último no ha sido debidamente explicitado[34].

Los problemas del determinismo son problemas referidos a la dependencia o no de fenómenos o eventos respecto de sus antecedentes: si se quiere se los puede formular en términos decimonónicos de cuerpo material, fuerza y movimiento, en un marco de mecánica clásica.
Los problemas de la libertad son problemas referidos a la dependencia o independencia de los actos de un ser con respecto al mundo exterior de ese ser. Aquí, siguiendo a Vaz Ferreira, utilizamos las palabras ser y acto como se las emplea y comprende en el lenguaje común.

---

[33] Después de la aparición y consolidación de las físicas relativista y cuántica, el concepto científico de materia propio del siglo XX tiene poco que ver con el del siglo XIX: es mucho menos rudimentario, pertenece a un nivel de abstracción más profundo y tanto él como los términos relacionados (como el de masa) han sufrido procesos de diferenciación de conceptos (aparición de los conceptos de masa de reposo y de materia cuya masa de reposo es nula, como es el caso de los fotones) y de integración de conceptos (confluencia o equivalencia de masa y energía, e intercovertibilidad de materia con masa de reposo y materia sin ella, como en la aniquilación electrón-positrón para dar origen a fotones o la creación de pares electrón-positrón a partir de fotones) muy significativos. Pero sigue habiendo un cierto grado de acuerdo tácito sobre cómo debe utilizarse ese término y los términos con él relacionados, ahora con sus nuevos significados, en el marco de las actuales teorías físicas.

[34] Si los términos materia y masa se utilizan en un análisis filosófico con el mismo sentido que se les asignaba en la física decimonónica, entonces el enunciado “toda materia posee masa” es verdadero. Pero si se les asigna el significado que tienen en la teoría restringida de la relatividad, y si se sobrentiende que con “masa” se alude a lo que en física relativista se denomina “masa de reposo”, entonces ese enunciado es falso, porque los campos electromagnéticos no poseen masa en reposo pero no obstante hay acuerdo en considerarlos parte de la materia.

Entonces, si un ser no es totalmente dependiente de lo que no es él, puede ser libre aunque sus actos en la medida que involucren eventos internos y externos al ser, puedan depender de todos sus antecedentes y estar así en conjunto, determinados.

La analogía del barco a vapor, que pone Vaz Ferreira, ilustra lo que se quiere decir.
El barco representa al ser, el mar y la atmósfera que lo rodean representan el mundo exterior al ser. Cuando la caldera no produce vapor y la hélice se encuentra detenida, las corrientes marinas y los vientos determinan el rumbo del barco: no es libre porque sus actos están determinados por su mundo exterior. Cuando la hélice se encuentra en funcionamiento su rumbo depende del impulso que suministre al barco combinado con el arrastre de las corrientes y los vientos: es parcialmente libre porque sus actos no están determinados enteramente por el mundo exterior.
Pero en ambas situaciones todos los eventos que involucran al barco están determinados por sus antecedentes.

Así pues, puede darse una respuesta positiva a un problema de libertad relativo a un ser aun cuando el correspondiente problema de determinismo relativo a los actos de ese ser tenga también una respuesta positiva. También puede concebirse una respuesta positiva al problema de la libertad junto con una respuesta negativa al problema del determinismo.
Pero, contrariamente a lo que a menudo se piensa cuando no se distingue entre los problemas de la libertad y los problemas del determinismo, para poder admitir el libre albedrío no es necesaria la indeterminación en el mundo físico introducida por la física cuántica.

Ernst Cassirer señalaba, al analizar el determinismo y el indeterminismo en la mecánica cuántica, que el libre albedrío de los seres humanos se vincula con la posibilidad de una conducta ética. Esta última está por lo general determinada por valoraciones y motivaciones que van de lo más abyecto a lo más sublime. La conducta ética parece ser en lo fundamental predecible, determinada, en vez de ser un fruto del azar. Desde el punto de vista psicológico los fenómenos relacionados con la voluntad resultan a menudo inescrutables desde el exterior del ser, pero aparecen en gran medida determinados cuando se los considera teniendo en cuenta al ser en el mundo. No son precisamente los huecos en el determinismo los que en los seres humanos se llenan con libre albedrío.
Como decía Bertrand Russell[35], "la libertad, en cualquier sentido que se la tome digno de tenerse en cuenta, solo pide que nuestras voliciones sean, como son, fruto de nuestros propios deseos y no efecto de una fuerza exterior que nos coaccione a querer lo que no quisiéramos…el libre albedrío es real en la única forma que importa."

Además del determinismo científico aplicado a los acontecimientos del mundo real, que es en principio refutable, se tiene otra modalidad, el **determinismo metafísico**.
Los que adoptan esta postura sostienen que todos los acontecimientos están predeterminados, que el futuro es tan inamovible como el pasado, con independencia de la existencia de observadores y de todo intento de predicción. Los sucesos están predeterminados inclusive aunque el mundo no manifestara regularidades y pareciera aleatorio.

[35] B. Russell "*Our knowledge of the external world*", Routledge, Londres, 1993, p.236.

Esta versión débil del determinismo no parece poseer suficiente contenido empírico como para ser contrastada y eventualmente refutada, pero para algunas personas es una convicción metafísica demasiado arraigada como para abandonarla.

3. Campos y principio de contigüidad

"Las teorías de campos…se formularon por primera vez en la mecánica de los medios continuos: acústica, hidrodinámica, teoría de la elasticidad. Allí, efectivamente, no era importante si la materia era concebida en principio como un continuo o como teniendo una estructura atómica. En el último caso, el tratamiento como continuo era desde el punto de vista empírico una buena aproximación, y así es como interpretamos hoy estas teorías…La creencia en una mecánica de masas puntuales con acciones a distancia permaneció como una fase transitoria de la física. Aún desde el análisis de la electrodinámica por Faraday y su formulación matemática por Maxwell, las fuerzas mismas fueron consideradas una realidad física con una dinámica interna, es decir, un campo. Por vez primera el espacio y el tiempo se describieron como un continuo tetra-dimensional y la dinámica se describió por medio de un sistema de ecuaciones a derivadas parciales hiperbólicas."
Carl von Weizsäcker, Thomas Görnitz y Holger Lyre en "The structure of Physics", Springer, Dodrecht, Holanda, 2006.

Además de cuerpos materiales, en el espacio físico hallamos campos.
A diferencia de los cuerpos, en general los campos no se pueden utilizar para definir marcos de referencia de espacio y tiempo. Los campos, también a diferencia de los cuerpos materiales, no tienen efectos directos sobre nuestros sentidos, son intangibles: sus propiedades se estudian a partir del comportamiento de los cuerpos, tangibles y ponderables, sobre los cuales se presume que actúan esos campos.
Se conviene en que dos campos son físicamente idénticos si su efecto sobre el mismo cuerpo de prueba es el mismo en cada punto del espacio donde se ubique dicho cuerpo. Por tanto los campos, a diferencia de las partículas, no se encuentran localizados sino que se encuentran distribuidos en el espacio. Los campos se detectan y se pueden estudiar midiendo las fuerzas que esos campos producen en cuerpos de prueba, o sea, en última instancia determinando sus efectos sobre el movimiento de los cuerpos. Esos cuerpos de prueba se introducen en la región ocupada por el campo con el fin de estudiarlo y por lo menos en principio, poder determinarlo en cada punto de esa región. Por este motivo las características de los cuerpos de prueba deben ser tales que la introducción de uno de esos cuerpos no debe modificar significativamente el campo.
Si se estudia un campo gravitatorio debido a un sistemas de masas, como cuerpo de prueba se debe utilizar uno cuya masa sea lo suficientemente pequeña como para no alterar la evolución en el tiempo y en el espacio de la configuración de los cuerpos que se consideran la fuente del campo, puesto que si así no fuera, dicha alteración alteraría el campo que se desea medir. Por ejemplo, un satélite artificial es un buen cuerpo de prueba para estudiar el campo gravitatorio debido al sol, los planetas y la luna en una región relativamente próxima a la tierra, pero un objeto cuya masa fuera similar a la de la luna no lo sería.
Lo mismo ocurre en relación con el campo eléctrico estático, cuyas fuentes se encuentran en un sistema de cargas. La carga de prueba debe ser lo bastante pequeña como para que el campo que produce no altere la evolución de las cargas del sistema cuyo campo interesa estudiar.

Los **campos de fuerza** se pueden clasificar en **conservativos** y **no conservativos**, a partir del trabajo que el campo realiza sobre un cuerpo de prueba cuando este último se

desplaza sobre una curva orientada que conecta dos puntos cualesquiera y es lo bastante regular.
Para calcular el trabajo se consideran las componentes de la fuerza tangentes a la curva en cada punto de ésta. Se elige un conjunto finito de puntos sobre la curva considerada, incluyendo los puntos extremos inicial y final. Se aproxima la curva mediante la poligonal definida por esos puntos. Se multiplica la componente tangencial de la fuerza en el punto inicial de cada segmento de la poligonal por la longitud del segmento y se suma sobre todos los segmentos. A medida que se eligen más puntos y más próximos, las sumas correspondientes se aproximan a un número real que por definición es el trabajo que el campo de fuerzas realiza sobre el cuerpo de prueba cuando éste recorre la curva.
Cuando el trabajo realizado es el mismo sobre todas las curvas orientadas que conectan un punto inicial con otro final cualesquiera, el campo se denomina conservativo. En caso contrario es no conservativo.
Cuando el campo de fuerzas es conservativo posee un potencial, es decir, se puede hallar una función escalar definida en cada punto del espacio tal que a partir de ella se pueden calcular las fuerzas mediante un operador vectorial adecuado (el operador gradiente). Y si un campo de fuerza posee un potencial, es conservativo.

3.1 Campos y energía en mecánica newtoniana de partículas

El concepto de campo aparece ya en la mecánica clásica, newtoniana, en la que se usa ese concepto para describir la fuerza de interacción entre partículas.
Se postula que cada partícula crea un campo alrededor de sí misma, campo que es capaz de ejercer una fuerza sobre otra partícula que allí se encuentre.
Cuando el campo de fuerzas que actúa sobre una partícula dada es conservativo, la energía definida como la suma de la energía cinética y el potencial (denominado energía potencial en este caso) se mantiene constante. Este es el **principio de conservación de la energía.**
La ley de atracción gravitatoria de Newton puede interpretarse en términos de campos de fuerza. Consideremos una masa M que se puede considerar puntual y ubiquemos en ella el origen de coordenadas a partir del cual medimos las distancias a los puntos del espacio. Si m es una masa de prueba situada a una distancia r del origen, y si $\vec{e}_r$ es un vector de longitud unidad orientado radialmente alejándose del origen, la fuerza de atracción sobre la masa de prueba viene dada por:

$$\vec{F}(r) = -G \cdot \frac{m_g \cdot M_g}{r^2} \cdot \vec{e}_r = m_g \cdot \left(-G \cdot \frac{M_g}{r^2} \cdot \vec{e}_r \right) = m_g \cdot \vec{g}(r)$$

El campo $\vec{g}(r)$ posee simetría esférica: está orientado radialmente, dirigido hacia el origen, y su magnitud es la misma en todos los puntos de la superficie de una esfera de radio r. En este caso el operador gradiente se reduce a $-\frac{d}{dr}(\ ) \cdot \vec{e}_r$.

Como $\vec{g}(r) = -\frac{d}{dr}\left(-\frac{G_g \cdot M}{r} \right) \cdot \vec{e}_r = -\frac{d\varphi(r)}{dr} \cdot \vec{e}_r$, resulta que $\vec{g}(r)$ posee un potencial $\varphi(r) = -G \cdot \frac{M_g}{r}$ y por tanto es un campo conservativo.

La segunda ley de Newton $m \cdot \vec{a} = \vec{F}$ para el movimiento de la masa de prueba adopta una forma independiente de dicha masa: $\vec{a}(r) = -\vec{g}(r)$

Es decir $m \cdot \frac{d\vec{v}}{dt} = m \cdot \frac{d^2\vec{r}}{dt^2} = m_g \cdot \left(-\frac{d}{dr}\varphi(r) \cdot \vec{e}_r \right)$

El teorema de conservación de la energía para la partícula de masa m moviéndose con velocidad v en el campo de potencial gravitatorio $\varphi(r)$ creado por la masa M adopta la forma $\frac{1}{2} \cdot m \cdot v^2 + \Phi(r) = E$

En esta expresión $\Phi(r) = m_g \cdot \varphi(r)$ es la energía potencial gravitatoria de la masa y E es su energía, que permanece constante.

La conservación de la energía puede extenderse a un sistema de partículas sometidas a fuerzas conservativas externas y de interacción entre las partículas del sistema.

La suma de las energías cinéticas de las partículas del sistema y la energía potencial $U(\{\vec{r}_j\})$ del mismo permanece constante[36]: $\frac{1}{2}\sum_j m_j \cdot v_j^2 + U(\{\vec{r}_j\}) = E$

El movimiento de un cuerpo relativo a un marco no inercial Ñ se describe, como vimos en el capítulo precedente, mediante una modificación de la segunda ley de Newton: $m \cdot \vec{a}' = \vec{F} - m \cdot \vec{W}$ La fuerza de inercia $-m \cdot \vec{W}$, al estar definida para cada punto del espacio, puede interpretarse como un campo de fuerzas proporcional a la masa inercial del cuerpo cuyo movimiento se estudia.

Cuando $\vec{W}$ resulta independiente del tiempo y de la posición, como ocurre cuando Ñ se mueve con movimiento rectilíneo uniformemente acelerado respecto del marco inercial K, el campo de fuerzas de inercia es análogo al campo de fuerzas gravitatorio en una región lo bastante pequeña como para que $\vec{g}(r)$ se pueda considerar a su vez independiente de la posición. Esta analogía resulta tanto más sorprendente cuando se tiene en cuenta la equivalencia de la masa inercial y la masa gravitatoria, $m_g = m$ La reflexión sobre este hecho conduce, según se verá posteriormente, a un nuevo punto de vista sobre la gravitación: el de la teoría generalizada de la relatividad.

3.2 Campos en mecánica de medios continuos y electrodinámica: la versión fuerte del principio de contigüidad.

No obstante las consideraciones sobre los campos efectuadas previamente en esta sección, tanto en mecánica clásica de partículas como en mecánica de cuerpos rígidos, se trabaja en condiciones en las que se puede considerar instantánea la propagación de las perturbaciones. En este caso el concepto de campo es conveniente pero posiblemente no es indispensable para estudiar el movimiento de un sistema de partículas o de un cuerpo rígido.

Muy distinta es la situación en mecánica de medios continuos, en ramas como la teoría de la elasticidad dinámica o la acústica. El estudio de los fenómenos propios de estas disciplinas no puede llevarse a cabo si no se tiene en cuenta el retardo en la propagación de las perturbaciones en los estados de esfuerzo y deformación locales. Otro tanto ocurre en la teoría electromagnética en condiciones en las que corrientes transitorias de

[36] $\{\vec{r}_j\}$ representa la configuración del sistema, dada mediante el conjunto de los vectores de posición de las partículas que lo componen, en cada instante de tiempo.

muy corta duración o corrientes alternas de frecuencias muy elevadas radian una fracción considerable de la energía del campo hacia el espacio circundante.

Cuando se tiene en cuenta la velocidad finita de propagación de las interacciones, cualquier modificación en una partícula se hará sentir siempre un tiempo después sobre cualquier otra, de modo que la fuerza de interacción entre dos partículas no puede depender solamente de sus posiciones en un mismo instante.

Si se parte del **principio de contigüidad**, admitiendo que la interacción es posible solamente entre puntos vecinos en el espacio, es más natural introducir un campo como portador de las modificaciones que se producen en las partículas.
El campo puede ser considerado una entidad real, en el siguiente sentido: una modificación en una partícula modifica el campo en su vecindad, esta modificación del campo se propaga (no necesariamente tan rápido como la luz en el vacío, pero si aceptamos la teoría de la relatividad, ciertamente no más rápido que eso) hasta que alcanza a la otra partícula, influyendo sobre ella. Y la modificación que esta última sufre se traslada al campo, se propaga a su vez en el campo y alcanza a la primera partícula. Según esta forma de ver las cosas, las partículas interactúan directamente con los campos, no con otras partículas.
Cuando las interacciones entre cuerpos son mediadas por campos cuyas modificaciones se propagan con velocidades finitas, cabe esperar que tanto la tercera ley de Newton[37] (igualdad de acción y reacción) como su ley de gravitación universal no sean válidas. Así estas leyes pierden su validez en el marco de la mecánica relativista. Sin embargo, son muy buenas aproximaciones dentro de los límites de validez de la mecánica newtoniana.

Con el desarrollo del concepto de campo, el principio de causalidad clásico tomó la forma dada por el principio de contigüidad. También se lo conoce como principio de acción cerrada, **de localidad**, o de ausencia de acción a distancia.

En forma precisa, en la **formulación fuerte** del principio se supone que los eventos en un punto P en un instante t depende solamente de los eventos en los puntos de una cierta región, próxima al punto P y que lo contiene en su interior, en un instante previo $t-\Delta t$. A medida que Δt decrece, las dimensiones de esa "región de influencia" sobre P disminuyen, tendiendo a cero cuando Δt tiende a cero. Si, como se hace en la teoría de la relatividad, se admite que la velocidad de propagación de las interacciones o las señales se encuentra acotada superiormente (por la velocidad c de la luz en el vacío) el estado en un punto P depende a lo sumo del estado de los puntos en una esfera de radio $c\cdot\Delta t$ en un instante $t-\Delta t$ previo.

3.3 Campos en la teoría de los procesos de transporte de calor y masa: la versión débil del principio de contigüidad.

Las ecuaciones de la electrodinámica, la elastodinámica y la acústica son compatibles con la formulación fuerte del principio de acción local.

[37] Cuando dos cargas interactúan a través del campo electromagnético, un cambio de posición en una de ellas se traduce en una variación de la fuerza que actúa sobre la otra en un instante posterior. Por este motivo no cabe esperar que en un instante dado las fuerzas de interacción entre las partículas sean de igual magnitud y dirección, y sentido opuesto, como debería ocurrir de acuerdo con la tercera ley.

Por el contrario, las ecuaciones del transporte de masa y calor (de tipo difusivo) violan la formulación fuerte del principio. Esto se debe a que las soluciones de estas ecuaciones difusivas implican que el efecto de una perturbación localizada en la concentración o en la temperatura, ocurrida en un instante dado, en un instante posterior se habrá propagado a todos los puntos del espacio, no importa cuán alejados se encuentren.
Por ejemplo, la ecuación del transporte de masa por difusión, en una dimensión espacial puede escribirse en términos del campo de concentraciones $c(t,x)$ y del coeficiente de difusión D de cierta sustancia, así: $\frac{\partial c}{\partial t} = D \cdot \frac{\partial^2 c}{\partial x^2}$
Supongamos que para $t < 0$ la concentración es nula en todas partes en un medio no acotado cuyas abscisas verifican $-\infty < x < +\infty$.
Si se introduce en forma instantánea, en el instante $t = 0$ y en el punto $x = 0$, una masa M de sustancia, entonces el campo de concentraciones para todo instante $t > 0$ y para todo punto de abscisa x viene dado por[38]:

$$c(t,x) = \frac{M}{2\sqrt{\pi \cdot D \cdot t}} \cdot \exp\left[-\frac{x^2}{4\pi \cdot D \cdot t}\right]$$

De esta solución se desprende que por pequeño que se tome $t > 0$ y por grande que sea x en valor absoluto, la concentración correspondiente a ese instante y en ese punto va a ser positiva.
En consecuencia las soluciones de esta clase de ecuaciones conllevan una velocidad infinita de propagación de las perturbaciones en la concentración o en la temperatura del medio. Pero como el efecto de la perturbación decrece rápidamente con el aumento de la distancia entre la región perturbada y el punto donde se considera el efecto, en la práctica éste no puede detectarse fuera de una cierta región de influencia finita que depende entonces de la apreciación de los instrumentos de medición.

3.4 El postulado de posibilidad de descripción continua de los fenómenos físicos

En cualquiera de sus versiones el principio de contigüidad presupone la posibilidad de una descripción continua en el tiempo y en el espacio. Tanto la física clásica como la física relativista dan por sentado este postulado de continuidad en la descripción.
Si se pone en duda su validez, se debe poner en duda el principio de contigüidad, una de las bases tanto de la teoría restringida como de la teoría generalizada de la relatividad.
Si se descarta la posibilidad de una descripción ininterrumpida del mundo físico, entonces la formulación más precisa del principio de causalidad en física, es decir, el principio de contigüidad, deja de ser universalmente aplicable.

Cuando medimos una cantidad física, obtenemos un conjunto finito de valores numéricos aproximados. Si buscamos relacionar los valores que toman dos cantidades, como la deformación de una varilla sometida a cargas de tracción diferentes, obtenemos un conjunto finito de pares de valores (deformación, carga).

[38] Esta solución se puede hallar en cualquier libro sobre difusión o transporte de calor, así como en los libros de matemática que dedican al menos una parte al estudio de las ecuaciones en derivadas parciales de tipo parabólico. En el libro clásico de H. Carslaw y J. Jaeger “Conduction of heat in solids”, Clarendon Press, Oxford, 1976, se puede encontrar un tratamiento exhaustivo del tema.

Si medimos la deformación para un conjunto de cargas más numeroso y con valores numéricos más próximos entre sí, obtendremos un conjunto de deformaciones correspondiente, más numeroso y con valores numéricos más próximos entre sí.
Al final, la apreciación de los instrumentos de medida pone un límite a este proceso: desde el punto de vista experimental no hay nada más.
En algunos casos se puede mejorar la apreciación y en consecuencia se puede obtener un conjunto de resultados más numeroso y denso. Pero suponer que la apreciación de los instrumentos se puede hacer tan pequeña como se desee no es algo que se pueda sustentar ni en evidencia experimental ni en resultados teóricos. Puede ser una suposición muy conveniente para elaborar un modelo, pero ahora se sabe que es falsa[39].

Siempre podemos representar las dependencias físicas tales como la relación entre la carga mecánica y la deformación de una varilla por medio del tipo de funciones conocidas como funciones analíticas.
Esta representación presupone que tanto la variable que corresponde a la carga como la variable que corresponde a la deformación pueden tomar cualquier valor real en un continuo numérico, y además requiere un grado muy elevado de regularidad en la relación entre las variables, cosa que escapa a toda experiencia posible.
Como señala Scrödinger[40], "suponer que la dependencia física **es** de este tipo simple, es un paso epistemológico desmesurado y seguramente inaceptable."

No obstante las soluciones de las ecuaciones diferenciales ordinarias que expresan en forma matemática las leyes de la mecánica newtoniana de partículas y las soluciones de las ecuaciones a derivadas parciales que expresan en forma matemática las leyes de los campos de la mecánica de medios continuos y de la electrodinámica, poseen por lo general ese grado elevado de regularidad. La dependencia regular entre las variables físicas es una consecuencia de la forma matemática de las leyes.
Esta forma matemática surge de conjeturas fundadas y es válida dentro de ciertos límites. Por su origen y por su grado de confirmación, cabe esperar que constituyan modelos adecuados del mundo medio, y efectivamente lo son. Pero en principio no deberíamos esperar que se puedan aplicar sin más al ámbito de lo microscópico o a escala de las galaxias.

Además, el mismo mundo medio presenta aspectos que requieren incorporar un grado de irregularidad en el modelado matemático de las leyes que las ecuaciones de la física-matemática clásica y las figuras inspiradas en la geometrías de Euclides no parecen estar en condiciones de aportar.
La descripción de las nubes, las cadenas de montañas, las costas, la estructura de los vórtices durante el movimiento turbulento del agua en ríos y mares, la estructura de la turbulencia de la atmósfera, y los relámpagos, no se puede realizar mediante aproximaciones a partir de esferas, conos, rectas y círculos.
De ahí la invención en 1975 (por Mandelbrot) y el exuberante desarrollo a partir de ese momento, de la geometría fractal y de temas matemáticos estrechamente relacionados

[39] La experiencia y la teoría sugieren que más tarde o más temprano se alcanza un punto en el que el nivel de ruido debido a la dinámica microscópica de la materia impide todo progreso en la mejora de la apreciación.

[40] E. Schrödinger, "Ciencia y humanismo", Tusquets, Barcelona, 2ª edición, 1998, p.p. 44-45.

cuyos antecedentes habían sido considerados como curiosidades sin interés para la descripción del mundo físico.
Una vez que se dispuso de ordenadores adecuados, el cálculo numérico de las soluciones de las ecuaciones diferenciales no lineales del tipo que puede hallarse en algunos modelos matemáticos de origen mecánico, puso en evidencia que esas soluciones presentan a veces un comportamiento asintótico[41] que parece aleatorio. El estudio de estas propiedades originó una rama de la teoría de sistemas dinámicos, la dinámica caótica, que se vincula naturalmente con la geometría de los fractales a través de los denominados atractores extraños[42].

La interpretación generalmente admitida de la mecánica cuántica descarta la posibilidad de una descripción ininterrumpida a nivel microscópico, eliminando de hecho el postulado de continuidad (que es en última instancia un postulado metafísico) y algunas de sus consecuencias, como el concepto de trayectoria de una partícula y su identidad. Puesto que los resultados de las mediciones y las relaciones determinadas experimentalmente entre las variables de un sistema son siempre conjuntos finitos, no hay nada en esto que por sí mismo impida eliminar el postulado de continuidad.
No obstante, la descripción del estado y de la evolución de un sistema cuántico de N partículas se puede realizar a partir de la denominada función de estado.
En la representación que emplea el tiempo y un espacio de configuraciones construido a partir de las coordenadas espaciales $\vec{r}_1,...,\vec{r}_N$ de las partículas, esa función de estado (denominada ahora función de ondas) $\Psi(t,\vec{r}_1,...,\vec{r}_N)$ es una función en general con valores complejos, de 3.N+1 variables reales (cada vector de posición de una partícula aporta tres coordenadas escalares, de modo que el espacio de configuración posee 3.N dimensiones).
Verifica una ecuación de campo que presupone no ya el postulado de descripción continua sino la versión fuerte del principio de contigüidad: la ecuación de Schrödinger.
Si conocemos $\Psi(t_0,\vec{r}_1,...,\vec{r}_N)$ para un instante inicial t_0 y las interacciones dentro del sistema y entre el sistema y su ambiente, entonces el campo a valores complejos $\Psi(t,\vec{r}_1,...,\vec{r}_N)$ queda determinado para todo tiempo y para toda configuración del sistema de partículas. Como el dominio de Ψ es un espacio multidimensional cuya dimensión puede ser muy elevada (si el sistema consta de 5 partículas, su dimensión es

[41] Es decir, cuando el tiempo aumenta a partir de un instante inicial y toma valores lo suficientemente grandes. Puesto que un ordenador es una máquina con un conjunto finito de estados internos y puede cambiar de estado en un conjunto discreto y finito de instantes de tiempo, una especificación rigurosa de la conexión entre la ecuación diferencial y el modelo discreto para el cálculo numérico y la representación gráfica de las soluciones de esa ecuación, es un asunto bastante más complicado de lo que parece a primera vista. Cuando se tiene una dinámica caótica las dificultades aumentan.

[42] En la teoría de sistemas dinámicos se representa el estado del sistema mediante un punto perteneciente al denominado espacio de estados del sistema. El espacio de estados es un espacio abstracto con una vecindad definida para cada punto. En las aplicaciones de los sistemas dinámicos en ciencias de la naturaleza se emplea una clase de espacios abstractos conocida como espacios métricos. En un espacio métrico la vecindad se define a partir de una distancia definida entre cada par de puntos del espacio. El conjunto de puntos por los que pasa el sistema durante su evolución se denomina órbita, y la correspondencia entre instantes de tiempo y puntos de la órbita se conoce como trayectoria. Un atractor es un conjunto de puntos en el espacio abstracto de estados del sistema dinámico al cual se aproximan las órbitas que comienzan a partir de puntos lo bastante próximos al atractor (en la denominada cuenca de atracción). Un atractor extraño se puede describir como fractal.

igual a 15, si consta de 100 partículas la dimensión es de 300) se ve que la función de ondas no puede interpretarse como un campo que ocupa el espacio físico, tal como el campo electromagnético, el campo gravitatorio u otros campos materiales.

Pero el cuadrado del módulo $\left|\Psi(t,\vec{r}_1,...,\vec{r}_N)\right|^2$ posee un significado físico bien determinado: es la densidad de probabilidad de encontrar el sistema, en el instante t, en la configuración $\vec{r}_1,...,\vec{r}_N$. En principio el conocimiento de la función de onda permite calcular la probabilidad de los resultados de la medida de una propiedad del sistema, aplicando el operador correspondiente. Un análisis detallado del proceso de medición mecánico-cuántico, empleando la interpretación más admitida tanto del formalismo matemático como de los resultados de los experimentos, conduce a la conclusión de que una descripción ininterrumpida no es posible: en el micro-mundo no existe algo que pueda equipararse con la trayectoria de una partícula en el meso y en el macro-mundo de las mecánicas newtoniana y relativista[43].

3.5 Reconsideración del principio de causalidad

A la luz de la mecánica cuántica la afirmación "todo evento tiene alguna causa" debe refinarse para tener en cuenta las leyes probabilísticas.

Se puede decir ahora "el evento concomitante o previo C en la cosa (o parte de una cosa) α, causa el evento E en otra cosa (o en parte de la misma cosa) β, cuando y solo cuando C genera una transferencia de energía de α hacia β, de la que resulta E con cierta probabilidad condicionada $p = P(E/C)$"

Esto flexibiliza el enunciado del principio de causalidad hasta el punto en que resulta posible que una causa no siempre produzca el efecto esperado.

Si la probabilidad condicionada es igual a 1, parecería que este enunciado probabilístico se reduce al determinista clásico. Pero si se admite que un evento con probabilidad 1 puede en teoría no ocurrir, esta interpretación del principio determinista como caso límite del principio probabilista no es aceptable sin un análisis a fondo del concepto de probabilidad a que hace referencia y su conexión con los resultados de la experiencia.

Otra observación pertinente llegado este punto en el desarrollo del tema, es que si los procesos causales se producen entre cosas diferentes o entre partes diferentes de una misma cosa, entonces los eventos no suscitados por influencias externas, los eventos que consideramos espontáneos, no tendrían causa identificable.

Por ejemplo, la desintegración radioactiva espontánea de algunos núcleos atómicos, o lo que podríamos denominar el "auto-movimiento" de un corpúsculo una vez eyectado desde un núcleo atómico debido a un evento de desintegración. No obstante, si se pueden hallar dos partes distintas en condiciones de intercambiar energía, un evento aparentemente espontáneo podría en realidad ser explicado por la versión probabilista del principio causal. Este es el caso de los dos ejemplos que pusimos: la desintegración

[43] Un tratamiento actualizado y detallado de la mecánica cuántica y sus problemas, con una exposición cuidadosa de base filosófica y accesible a un amplio círculo de lectores, puede hallarse en el libro de M. Guerra, "El enigma cuántico", coeditado por Grupo Magro y la Universidad Católica del Uruguay, Montevideo, 2009.

radioactiva (por efecto túnel mecánico-cuántico o por formación de un racimo lo suficientemente estable y energético de nucleones como para salir del núcleo en forma clásica) y el auto-movimiento del corpúsculo eyectado del núcleo de alguna de las maneras recién mencionadas.
Resignarse a eventos no causados no parece algo comprendido dentro del horizonte de la ciencia.

4. La electrodinámica clásica

"Es difícil hallar una teoría de campos en física clásica o en física moderna que haya tenido un impacto comparable al de la teoría del electromagnetismo y que, al mismo tiempo, tenga una reputación de casi incuestionable validez general. Parece seguro afirmar que desde el punto de vista macro-físico, la teoría del electromagnetismo es correcta y confiable, y que es útil y adecuada en numerosos ámbitos micro-físicos. Los enunciados precedentes no deben interpretarse en el sentido de que la teoría puede suministrar todas las respuestas en física y en tecnología dentro de los confines de las aplicaciones macroscópicas. Más bien indican confianza en la validez general de las ecuaciones de Maxwell. No obstante, todo problema concreto requiere especificar las propiedades electromagnéticas del medio bajo consideración, tal como se expresan a través de las relaciones constitutivas. Es una opinión generalizada que las discrepancias entre teoría y experimento se han debido al uso de relaciones constitutivas inadecuadas más que a inadecuaciones de las ecuaciones de campo de Maxwell mismas."

E. Post, "Formal structure of electromagnetics", North-Holland, Amsterdam, 1962.

En este capítulo se resumen algunos aspectos de la electrodinámica clásica, la disciplina que aborda el estudio de los campos electromagnéticos y sus fuentes en las cargas, los dipolos magnéticos y las corrientes eléctricas.

Las acciones eléctricas y magnéticas se ejercen sobre y entre diferentes cuerpos materiales. Si se admite que no existe acción a distancia, parece razonable suponer, como vimos en el capítulo precedente, que cuando un cuerpo actúa sobre otro debe existir un medio que los une y a través del cual se transmiten las acciones. Si se asume esto, **las acciones son siempre acciones de contacto**. Pero como las acciones eléctricas y magnéticas subsisten en el vacío, parecería que la existencia de fenómenos electromagnéticos implica, además de una modificación del estado de los cuerpos materiales, una modificación del estado físico del vacío. Esto conduce a introducir el concepto de campo como objeto perteneciente al mundo real. El campo deja de ser un mero concepto auxiliar, conveniente desde la perspectiva del análisis matemático de los fenómenos, como ocurre en la formulación original de la mecánica newtoniana de partículas, y pasa a tener el mismo status que la masa, la carga y la fuerza.

Este punto de vista se ve reforzado porque los experimentos sugieren que el equilibrio eléctrico no se establece en forma instantánea. Cuando la distancia entre los cuerpos es lo bastante grande y la duración de un transitorio eléctrico en uno de ellos es lo bastante pequeña, se observa que durante ese transitorio no aparece una modificación detectable en el estado eléctrico del otro cuerpo. Esta modificación aparece después de finalizado el transitorio.

Si admitimos que se puede y se debe efectuar una descripción continua de estos sucesos, parece ineludible suponer que entre las modificaciones en el estado eléctrico del primer cuerpo se transmite al vacío y a través de este último alcanzan posteriormente al segundo cuerpo (como ondas electromagnéticas viajeras).

El vacío, en la medida en que puede sufrir modificaciones físicas, es lo que los científicos del siglo XIX denominaron éter.

El campo electromagnético en una región del espacio vacío se describe dos vectores, un vector eléctrico y un vector magnético. Cada vector se describe por su magnitud, dirección y sentido. En cada instante esos dos vectores son asignados a cada punto de la región considerada, resultando así dos campos vectoriales.

En general la magnitud, dirección y sentido de estos vectores pueden variar de un punto a otro del espacio, y en el mismo punto, de un instante a otro.
La interacción de un campo electromagnético con una partícula depende de una cantidad escalar, la carga, que caracteriza la partícula a los fines de esta interacción, y de una cantidad vectorial, la velocidad de la partícula respecto de un marco de referencia inercial.
La carga, es sabido, puede ser positiva, negativa o cero. Esto permite distinguir partículas con carga, como los electrones y los protones, de partículas sin carga como los neutrones. La carga, tomada por separado, permite describir la interacción de la partícula con el campo eléctrico.
La velocidad, una característica cinemática, multiplicada por la carga, permite describir la interacción entre la partícula y el campo magnético.
El producto de la velocidad por la carga se puede relacionar con la corriente eléctrica. Esta última se mide determinando la cantidad de carga que atraviesa un área por unidad de tiempo, debido al movimiento de un conjunto de partículas.

4.1 Electro-estática, magneto-estática y electrodinámica pre-relativista

Consideremos ahora como surgen los campos eléctrico y magnético, comenzando con los campos estacionarios, o sea, que no varían con el paso del tiempo.
Comencemos con el hecho siguiente: las cargas eléctricas, positivas o negativas, producen campos eléctricos.
Las cargas no dependen del marco de referencia que se emplee para describir el movimiento: son invariantes relativistas[44]. Además, las cargas se conservan: si la carga total de un sistema de partículas se modifica, ello ocurre porque se transfiere carga desde el ambiente hacia el sistema o desde el sistema hacia el ambiente. La carga no se crea ni desaparece. Eso no obsta para que en ciertas condiciones una partícula cargada y su antipartícula, que posee la misma carga pero con signo opuesto, se aniquilen entre sí dando origen a fotones muy energéticos (radiación gama) que no poseen carga eléctrica: la carga total era cero antes de la aniquilación y continúa siendo cero después de la misma. La misma conservación de la carga ocurre cuando un fotón gama se transforma en un par formado por electrón negativo y por su antipartícula, el electrón positivo o positrón.
Los cuerpos con carga eléctrica producen fuerzas sobre cargas de prueba situadas en su cercanía. La fuerza de origen eléctrico F_E (un vector) sobre una carga de prueba depende de la magnitud de la carga q (un escalar) y de una propiedad local del espacio, el campo eléctrico E (un vector).
En notación vectorial se tiene: $\vec{F}_E = \mathrm{q} \cdot \vec{E}$
El campo eléctrico en condiciones estáticas es conservativo, por lo que se puede introducir una energía potencial y un potencial escalar. La fuerza de interacción entre dos partículas con cargas q_1 y q_2 es análoga a la fuerza de atracción gravitatoria, con

[44] Con el propósito de completar rápidamente el panorama, se llama la atención sobre leyes muy generales, como la de conservación de la carga, que se mantiene válida aún cuando otras leyes, como la ley de conservación de la masa, se han tenido que abandonar una vez establecidas la física relativista y la física cuántica.

la diferencia que puede ser atractiva cuando las cargas poseen distinto signo y repulsiva cuando poseen el mismo signo[45]: $F_E \approx \frac{q_1 \cdot q_2}{r^2}$

A esta fuerza de interacción le corresponde una energía potencial $U_E \approx \frac{q_1 \cdot q_2}{r}$

Introduciendo los potenciales electroestáticos $\varphi_1 \approx \frac{q_2}{r}$ producido por la carga q_2 en el sitio donde se encuentra q_1 y $\varphi_2 \approx \frac{q_1}{r}$ producido por q_1 en el sitio donde se encuentra q_2, resulta $U_E = \frac{1}{2} \cdot q_1 \cdot \varphi_1 + \frac{1}{2} \cdot q_2 \cdot \varphi_2$

Esta fórmula se puede generalizar para un sistema con un número cualquiera de cargas:

$$U_E = \frac{1}{2} \cdot \sum_{\alpha} q_\alpha \cdot \varphi_\alpha$$

En esta fórmula φ_α es el potencial del campo producido por todas las demás cargas en el punto donde se encuentra q_α.

Descomponiendo un cuerpo cargado extenso en un conjunto finito de elementos de volumen, se puede emplear la fórmula precedente, pasando al límite cuando el número de elementos tiende a infinito y sus dimensiones tienden a cero, para calcular la energía potencial de ese cuerpo debida a su propia distribución de carga: su auto-energía electroestática. Si se trata de una esfera de radio R_0 y carga total q uniformemente distribuida, su auto-energía es proporcional a $\frac{q^2}{R_0}$. Este resultado será retomado en 8.2, junto con el concepto de partícula elemental que se desprende de la teoría de la relatividad, para discutir los límites de la aplicación de la electrodinámica clásica.

Los cuerpos magnetizados (imanes), en reposo respecto de un sistema de referencia, producen fuerzas sobre otros cuerpos magnetizados y sobre cargas eléctricas en movimiento.

No se ha podido hallar hasta el momento cargas magnéticas aisladas (los denominados monopolos magnéticos) de modo que el elemento más simple con propiedades magnéticas es el dipolo magnético. Una corriente eléctrica en forma de espira (si se quiere, circular) se comporta como un dipolo magnético cuando se estudia el campo que produce a una cierta distancia, pero al parecer no todo dipolo magnético consiste en una corriente: el espín es el ejemplo típico. Como consecuencia, no toda interacción entre imanes se puede reducir a una interacción entre corrientes.

La fuerza de origen magnético $\vec{F}_B$ (un vector) sobre una partícula de carga q que se mueve con velocidad $\vec{v}$ (un vector) depende del producto de la carga por la velocidad y

[45] Esta expresión para la fuerza electroestática, al igual que las expresiones para la energía potencial y el potencial electroestático son igualdades en el sistema gaussiano de unidades físicas.
En el sistema internacional se transforman en igualdades luego de multiplicar los segundos miembros por $\frac{1}{4 \cdot \pi \cdot \varepsilon_0}$ donde ε_0 es la permitividad del vacío.

de una propiedad local del espacio en la cercanía del cuerpo magnetizado, el campo magnético $\vec{B}$ (un vector).

En notación vectorial se tiene: $\vec{F}_B = q \cdot \left(\vec{v} \times \vec{B}\right)$

El símbolo $\times$ representa el producto vectorial, en este caso aplicado al par vectorial formado por la velocidad y la inducción magnética. Es un vector perpendicular tanto a la velocidad como al campo, posee una magnitud igual al producto de la magnitud de la velocidad, la magnitud del campo de inducción y el seno del ángulo entre la velocidad y el campo magnético y se orienta en el sentido de avance de un tirabuzón dextrógiro cuando se gira desde el primer vector hacia el segundo vector.

Las cargas en movimiento constituyen corrientes eléctricas. Las corrientes eléctricas producen campos magnéticos. La ley de Ampere relaciona el campo magnético con las corrientes eléctricas que lo producen.

Las corrientes eléctricas variables en el tiempo producen campos magnéticos variables.

A su vez, un campo magnético variable en el tiempo en una región del espacio, induce en esa región un campo eléctrico.

En igualdad de las demás condiciones, cuanto más rápida es la variación del campo magnético, tanto más intenso resulta el campo eléctrico inducido.

El campo eléctrico inducido por la variación del campo magnético no es conservativo: la fuerza electromotriz, es decir el trabajo realizado por el campo eléctrico inducido sobre una carga positiva de magnitud unidad a lo largo de una curva cerrada, en general no es nulo: es igual a menos la velocidad de variación del flujo de inducción magnética a través de una superficie limitada por la curva. El cálculo del flujo exige orientar la superficie según el sentido positivo asignado a la curva.

La forma cuantitativa precisa de esta relación viene dada por la ley de inducción de Faraday. Como consecuencia del signo de menos que aparece en la ley de Faraday, el campo magnético asociado con la corriente eléctrica inducida tiende a oponerse a la variación del campo de inducción que en última instancia le dio origen. Es posible demostrar que si así no fuera, se violaría la ley de conservación de la energía.

Por su parte, un campo eléctrico variable en el tiempo en una región, induce a su vez un campo magnético en esa región. También en este caso y en igualdad de las demás condiciones, cuanto más rápida es la variación del campo eléctrico, tanto más intenso resulta el campo magnético inducido por esa variación.

Maxwell generalizó la ley de Ampere: además de la conexión ya mencionada entre el campo magnético y el movimiento de partículas cargadas, añadió una conexión entre el campo magnético inducido y la velocidad de variación del campo eléctrico.

De esta generalización se desprende que aunque la corriente de partículas cargadas sea nula, la variación temporal del campo eléctrico por sí sola alcanza para generar un campo magnético.

A su vez, la variación temporal del campo magnético genera un campo eléctrico.

La formulación matemática precisa de estas relaciones entre los campos eléctrico y magnético, perteneciente a la denominada electrodinámica microscópica, muestra que es posible generar **ondas electromagnéticas** que se propagan en el vacío lejos de las fuentes del campo con velocidad $c = 2{,}99776 \times 10^{10}$ cm/s.

Cuando el sistema de cargas microscópicas es muy complejo se puede enfocar el electromagnetismo desde un punto de vista macroscópico, menos detallado. El campo

electromagnético en una región del espacio ocupada por un medio material se describe entonces mediante un conjunto de cuatro campos (que de momento podemos considerar como si fueran campos vectoriales, aunque como veremos posteriormente en 7.5 en realidad frente a un cambio de marco de referencia inercial sus componentes no se comportan como las de un vector propiamente dicho): a saber dos campos eléctricos, y dos campos magnéticos. Estos cuatro campos se pueden introducir desde el comienzo o se pueden obtener mediante operaciones de tomar promedios sobre un sistema microscópico de partículas cargadas moviéndose en el vacío. Se obtiene así lo que se denomina electrodinámica de medios continuos o electrodinámica macroscópica. Nuevamente se deduce que en un medio continuo se pueden propagar ondas electromagnéticas. Si el medio continuo se puede considerar lineal, homogéneo e isótropo, como es el caso del aire, la descripción de las ondas se simplifica mucho y se puede realizar con dos campos, uno eléctrico y el otro magnético.

La confirmación experimental de la predicción sobre la propagación de ondas electromagnéticas en el aire, obtenida a partir de la teoría electromagnética, realizada por Hertz en el último cuarto del siglo 19, sentó las bases de la radiotecnia, la televisión y la técnica de micro-ondas tal como las conocemos en la actualidad.
La reducción de la teoría ondulatoria de la luz a la teoría de la propagación de ondas electromagnéticas permitió unificar la óptica con el electromagnetismo.

Las **fuentes del campo** propagado son corrientes eléctricas variables: en última instancia, **cargas eléctricas aceleradas**. Toda carga acelerada, de acuerdo con la electrodinámica basada en la teoría de Maxwell, radia ondas electromagnéticas en mayor o menor cuantía.
Si se trata de una partícula que transita en una órbita cerrada en torno de un centro de fuerza atractivo, como se suponía que era el caso de los electrones del átomo girando alrededor del núcleo atómico, la radiación tiene como consecuencia la disminución de la energía de la partícula, la modificación de su órbita y la aproximación progresiva al centro de fuerza.
Una estructura atómica semejante no puede ser estable: si se admite que toda carga acelerada radia, entonces la mecánica de las masas medias no puede ser aplicable al estudio de las estructuras atómicas y moleculares estables, propio de la Química.

4.2 La electrodinámica de los cuerpos en movimiento

Desde que Faraday introdujo las líneas de campo emanando de o arribando a una carga eléctrica o un polo magnético, se planteó la cuestión de determinar qué pasaba con esas líneas cuando las cargas o los polos se ponían en movimiento respecto del marco de referencia del laboratorio.
Para dar respuesta a estas interrogantes fue necesario esperar a disponer, primero de la teoría electromagnética de Maxwell, y segundo de una extensión de la teoría adecuada para abarcar el caso de cuerpos en movimiento. Una primera versión de esta extensión se debe a Hertz. Una segunda versión se debe a Lorentz. La tercera, y al menos por ahora última versión, es la debida a Einstein.
Dos generalizaciones empíricas se encuentran en la base de estas tres teorías sobre la electrodinámica de los cuerpos en movimiento:

(1) Los fenómenos de inducción (es decir, de génesis de fuerzas electromotrices inducidas) se producen por igual cuando varía el flujo magnético a través de un circuito de alambre conductor inmóvil respecto del laboratorio como cuando el campo magnético no varía respecto del laboratorio pero el circuito se mueve respecto del campo.
(2) Las fuerzas electromotrices inducidas dependen solamente de la velocidad relativa entre el circuito móvil y el campo fijo.
Solo la interpretación que se les da es diferente según qué teoría se considere.
Tanto en la formulación de Hertz como en la de Lorentz, se admite la existencia del éter.
La diferencia entre ambas se puede ver con mayor claridad recurriendo a la formulación de los campos de desplazamiento eléctrico $\vec{D}$ y de inducción magnética $\vec{B}$ en un cuerpo material, a saber[46]:

$$\vec{D} = \varepsilon_0 \cdot \vec{E} + \vec{P} \qquad \vec{B} = \mu_0 \cdot \vec{H} + \vec{M}$$

El campo $\vec{P}$ es la polarización dieléctrica (el momento dipolar eléctrico por unidad de volumen del medio material) mientras que el campo $\vec{M}$ es la magnetización (momento dipolar magnético por unidad de volumen del medio material).
Hertz asigna propiedades semejantes a los campos $\vec{E}$ y $\vec{P}$ por un lado, y a $\vec{H}$ y $\vec{M}$ por otro: ambos son arrastrados completamente por el cuerpo material en movimiento. Una dificultad que aparece como consecuencia de esta hipótesis es que las ondas electromagnéticas deberían ser arrastradas completamente por un gas en movimiento, no importando su grado de dilución. Los experimentos de Fizeau sobre propagación en medios en movimiento muestran que esto no ocurre: cuando el gas se encuentra lo bastante diluido, la propagación de las ondas electromagnéticas se produce con independencia del flujo del gas.
Para eludir esta dificultad Lorentz supone que los campos $\vec{E}$ y $\vec{H}$ permanecen fijos al éter, mientras que los campos $\vec{P}$ y $\vec{M}$, ligados a la materia, se mueven arrastrados por ésta. Pese al éxito de esta nueva formulación de la electrodinámica de los cuerpos en movimiento, los resultados negativos de los experimentos del tipo del de Michelson y Morley o de Trouton y Noble sugieren que el éter no existe.

En la formulación de Einstein se elimina el éter. Las ecuaciones que describen el campo electromagnético, las ecuaciones de Maxwell son compatibles con el principio de relatividad de Einstein si se supone que poseen la misma formulación matemática respecto de todos los marcos de referencia inerciales[47].
El campo electromagnético se considera ahora como una unidad y se describe respecto de cualquier marco de referencia inercial estableciendo cómo se transforman sus componentes al pasar de un marco a otro. Las leyes de transformación de los campos

---

[46] En el sistema internacional de unidades, donde ε_0 es la permitividad y μ_0 la permeabilidad del vacío.

[47] Lo cual, entre otras consecuencias, implica que la velocidad de la luz en el vacío posee el mismo valor numérico respecto de todos los marcos de referencia inerciales.

muestran que los campos eléctrico y magnético tomados por separado pueden ser nulos respecto de un marco de referencia inercial y no serlo respecto de otro marco distinto.
Puesto que el campo electromagnético actúa como un medidor de las interacciones entre los cuerpos materiales, constituye un sistema independiente al que pueden asignarse energía, cantidad de movimiento lineal y cantidad de movimiento angular como a cualquier otro sistema físico. Entonces, cuando el campo se encuentra aislado o cuando el campo y un conjunto de cargas que interactúan con él forman un sistema aislado, cabe esperar que se cumplan las correspondientes leyes de conservación para la energía total, para la cantidad de movimiento lineal total y para la cantidad de movimiento angular total del sistema. Los resultados de los experimentos sobre la presión de la radiación y el éxito en el diseño, construcción y operación de las máquinas eléctricas y los aceleradores de partículas, que involucran estos aspectos del campo electromagnético, confirman lo dicho anteriormente.

4.3 La teoría clásica del electrón

Durante la última década del siglo XIX se estudiaron las propiedades físicas de modelos de cuerpos materiales sólidos formados por electrones sometidos a fuerzas elásticas, de origen electroestático, centradas en un punto de equilibrio que se interpretaba como centro de las cargas positivas del átomo (las cargas positivas se consideraban distribuidas como un medio continuo en forma de esfera).
Una pregunta que surgía naturalmente era qué pasaba cuando ese ordenamiento de átomos y sus electrones se ponía en movimiento: cómo se modificarían las fuerzas y cómo esta modificación incidiría a su vez sobre las frecuencias de vibración de los electrones.
Como una carga en movimiento produce un campo eléctrico variable en un punto dado del espacio, un campo eléctrico variable genera un campo magnético, y un campo magnético en general ejerce una fuerza sobre una carga en movimiento, de todo esto se desprende que, cuando el sistema de cargas se pone en movimiento, aparecen nuevas fuerzas de interacción que se añaden a las electroestáticas.
A su vez, si dos polos magnéticos se encuentran en movimiento, pero permanecen en reposo uno respecto del otro, generan un campo magnético variable, el cual a su vez produce un campo eléctrico. Ahora bien, Lorentz mostró que un polo magnético en movimiento, en la presencia de un campo eléctrico, sufre una fuerza, en forma análoga a cómo una carga eléctrica en movimiento sufre una fuerza cuando se encuentra en un campo magnético. Como consecuencia, dos polos magnéticos que se desplazan solidarios ejercen una fuerza eléctrica el uno sobre el otro.
Al investigar las transformaciones de las cantidades mecánicas con el movimiento de un cuerpo formado por cargas eléctricas en interacción, Lorentz halló que las leyes de transformación parecían ser mucho más complicadas que las correspondientes a los campos. Debido a ello decidió profundizar en la dinámica de un electrón acelerado utilizando como modelo de electrón una porción de medio continuo con una carga eléctrica distribuida y que en reposo respecto del éter adoptaría la forma de una esfera de radio r_0 (denominado radio clásico del electrón).

Para explicar por qué el electrón no estallaba debido a las fuerzas repulsivas de los elementos de carga negativa que lo componían según este modelo, Lorentz, al igual que Abraham antes que él, introdujo en forma de hipótesis unas fuerzas atractivas de origen no eléctrico. Abraham había postulado que esas fuerzas de empaquetadura tomaban valores tales que el electrón mantenía su forma esférica cuando era acelerado respecto del éter. Lorentz se dio cuenta que este postulado conducía a contradicciones y dedujo que el único postulado razonable era exigir que como consecuencia de las fuerzas no eléctricas el electrón se encogiera en la dirección de su movimiento de modo que su dimensión longitudinal pasaría a ser $2 \cdot r_0 \cdot \sqrt{1 - \frac{v^2}{c^2}}$ en lugar de $2 \cdot r_0$, siendo v su velocidad respecto del éter. Las dimensiones del electrón transversales a su dirección de movimiento se conservarían independientes de la velocidad. Esta contracción había sido sugerida por FitzGerald, varios años antes, para poder explicar el resultado negativo del experimento de Michelson y Morley sin abandonar la idea del éter.

Una vez hallado el sistema de esfuerzos mecánicos internos que impedía la auto-destrucción del electrón, Lorentz encaró el cálculo de la fuerza que cada elemento de volumen cargado ejercía sobre cada uno de los otros elementos, en el caso de un electrón acelerado por un campo eléctrico externo[48].
En este cálculo de la auto-fuerza tuvo en cuenta que cuando el efecto de un elemento de carga se hace sentir sobre otro, el primero se ha desplazado a otra posición respecto del éter.
Conectó esta nueva posición con la posición previa mediante un desarrollo en serie de potencias del tiempo que las separa. Refirió el desarrollo en serie a la nueva posición, y lo expresó como suma de términos de posición, de velocidad de cambio de la posición (su aceleración) y de la velocidad de cambio de su aceleración, cortando el desarrollo a partir de este último término.
Integrando todas las fuerzas que unos elementos de carga ejercían sobre los demás, halló una expresión para la auto-fuerza total actuante sobre el electrón.
Introduciendo el desarrollo en serie de potencias mencionado en la expresión de la fuerza y efectuando algunas aproximaciones adicionales, Lorentz obtuvo una expresión consistente en la suma de dos términos. El primero, generalmente dominante, era proporcional a la aceleración a del electrón respecto del éter, la constante de proporcionalidad siendo negativa e igual a $-\frac{2}{3} \cdot \frac{e^2}{r_0 \cdot c^2}$ (en unidades gaussianas). El otro término, generalmente de orden inferior respecto del primero, era proporcional a la velocidad de cambio $\frac{d}{dt} a$ de esa aceleración.
Además, del cálculo se desprendía que la suma de la auto-fuerza y la fuerza $e \cdot E$ debida al campo eléctrico externo E era nula.
Dejando solamente el término dominante de la auto-fuerza y denominando $-m_0$ a la constante de proporcionalidad mencionada, la ecuación de equilibrio entre la auto-fuerza y la fuerza externa se podía re-escribir así: $m_0 \cdot a = e \cdot E$

[48] En el libro de D. Mentzel, "Mathematical Physics", Dover, New York, USA, 1961, pp. 317-325, se puede hallar un planteo completo y claro de la masa electromagnética y las fuerzas que actúan sobre una carga acelerada.

¡Pero ésta es una deducción de la segunda ley de Newton para el movimiento de un electrón, con la masa $m_0 \approx \frac{e^2}{r_0 \cdot c^2}$ expresada en términos del radio r_0 y la carga e del electrón, junto con la velocidad de la luz c en el éter!
El siguiente término en la expresión de la auto-fuerza, así como los términos de órdenes superiores que se obtienen si el desarrollo en serie no se trunca, miden la acumulación de energía electromagnética en el electrón a medida que la velocidad de éste aumenta.
Un análisis más detallado del movimiento permitió a Lorentz formular la segunda ley de Newton de esta forma: $\frac{d}{dt} p = e \cdot E$ siendo $p = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}} \cdot v$

Se advierte que la cantidad de movimiento p del electrón corresponde al de una partícula cuya masa m varía con su velocidad de acuerdo con la ley: $m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}$

Este resultado coincide con el que suministra la física relativista para una partícula cualquiera (no necesariamente un electrón) cuando m_0 se interpreta como la denominada masa en reposo.
Si bien la deducción electromagnética hecha por Lorentz de la masa inercial $m_0 \approx \frac{e^2}{r_0 \cdot c^2}$ y de la segunda ley de Newton resultaba muy atractiva, la evolución ulterior de la teoría física la alejó mucho de este punto de vista. El modelo del electrón como una porción de medio continuo parece demasiado simple. El electrón se comporta más como un elemento "puntual" que como un cuerpo extenso de un cierto radio r_0 y su masa no parece ser de origen electromagnético. Además, vista desde la perspectiva actual, la deducción de Lorentz es pasible de varias críticas[49].
Sea como fuere, la teoría de electrón de Lorentz es un antecedente valioso e inmediato de la teoría de la relatividad. Además, junto con los aportes de Poincaré al principio de relatividad, es una muestra de que a fines del siglo XIX y primeros años del siglo XX el ambiente intelectual estaba ya maduro para un cambio de paradigma.

[49] Ver, por ejemplo el libro de J.D.Jackson, "Classical Electrodynamics", 2º edición, Wiley, New York, USA, 1975, pp790-791.

5. Las ciencias físicas al final del siglo XIX

"En los cuarenta años previos a 1900, el electromagnetismo y la óptica fueron correlacionados y explicados en forma triunfal por la teoría ondulatoria basada en las ecuaciones de Maxwell. Puesto que la experiencia previa con los movimientos ondulatorios había involucrado siempre un medio para la propagación de las ondas, era natural que los físicos asumieran que la luz necesitaba un medio en el cual propagarse…
Las ondas sonoras y otros fenómenos ondulatorios similares son una consecuencia de la mecánica clásica: la existencia de marcos de referencia privilegiados respecto de los cuales los fenómenos son simples se comprende bien en términos de movimientos masivos del medio de propagación.
Pero para las perturbaciones electromagnéticas el medio parecía verdaderamente etéreo sin otra manifestación o propósito que soportar la propagación de las ondas. Cuando Einstein comenzó a pensar acerca de esos asuntos, existían varias posibilidades:

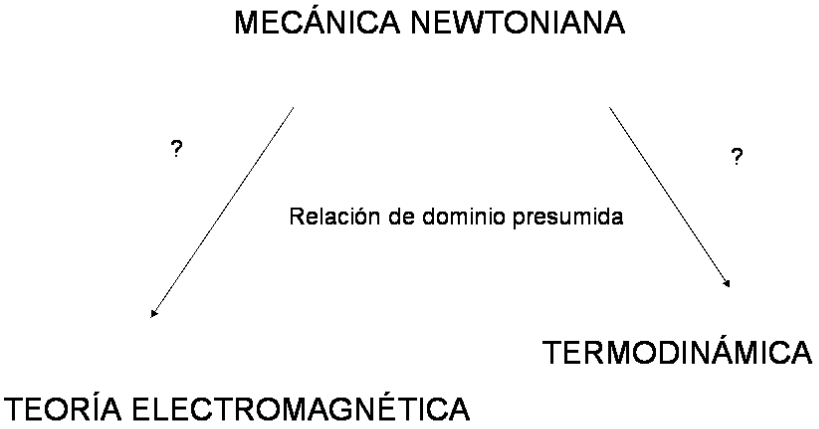
1. *Las ecuaciones de Maxwell eran incorrectas. La teoría adecuada del electromagnetismo debía ser invariante bajo las transformaciones de Galileo.*
2. *El principio de relatividad de Galileo se aplicaba a la mecánica clásica, pero el electromagnetismo poseía un marco privilegiado de referencia, respecto del cual el éter se encontraba en reposo.*
3. *Existía un principio de relatividad para ambas, mecánica clásica y teoría electromagnética, pero no era el principio de relatividad de Galileo. Esto implicaría la necesidad de modificar las leyes de la mecánica."*

J. Jackson "Classical electrodynamics", Wiley, New York, 1975.

Como dijimos en 1.5, el desarrollo de los fundamentos matemáticos de la mecánica clásica requirió y estimuló a su vez el desarrollo del cálculo infinitesimal.
La necesidad de contrastar sus predicciones con los resultados de las mediciones motivó la invención y puesta a punto de nuevos procedimientos de medición.
La combinación de las leyes mecánicas con métodos matemáticos poderosos y técnicas de medición refinadas mostró un inmenso poder explicativo.
Debido a todo esto, hacia 1800 la mecánica newtoniana de partículas se había convertido ya en el paradigma para la interpretación de todos los fenómenos físicos y para el desarrollo del aparato de la física-matemática, incluyendo el electromagnetismo.
Pero a finales del siglo XIX, a medida que se iban conociendo los resultados de nuevos experimentos, se hacía cada vez más difícil compatibilizar la mecánica newtoniana con la teoría electromagnética sobre la base del concepto de éter.

5.1 La mecánica newtoniana como súper-teoría física

Desde el punto de vista de su trabazón conceptual la física de la segunda mitad del siglo XIX se podía dividir en tres grandes teorías: mecánica, electrodinámica y termodinámica. Se admitía que la mecánica newtoniana debía considerarse como una especie de súper-teoría a la que eventualmente se lograría reducir las otras ramas de la física:



Las flechas de trazo lleno representan una relación de dominio presumida en esa época. Se suponía que la disciplina a la que llega la flecha debía ser dominada por la disciplina de donde parte la flecha.
Los signos de interrogación que aparecen al lado de las flechas expresan el carácter conjetural de esa suposición.
Esa relación de dominio debe entenderse como la capacidad de la mecánica newtoniana de dar cuenta tanto de la electrodinámica como de la termodinámica, en el sentido siguiente.
En esa época se suponía que los conceptos y leyes, tanto de la electrodinámica como de la termodinámica, se podrían reducir a los conceptos y leyes de la mecánica newtoniana, que las unificaría entonces en un mismo cuerpo de teoría física, basado en las nociones de espacio y tiempo absolutos.
Esta suposición se demostró errónea. La electrodinámica resultó ser compatible con la mecánica relativista, cuya formulación requirió la eliminación de los conceptos de espacio y tiempo absolutos, mientras que la termodinámica ha resultado ser compatible con, y ha podido reducirse en parte a la física cuántica.

5.2 El problema del éter

En 1880 la electrodinámica ya estaba formulada utilizando los conceptos de campo introducidos por Faraday y desarrollados matemáticamente como un sistema de ecuaciones diferenciales a derivadas parciales (ecuaciones de campo) por Maxwell. Los procesos de transporte de calor y masa también habían sido formulados utilizando ecuaciones de campo.
Pero mientras que las ecuaciones de la electrodinámica implicaban una velocidad finita de propagación de los efectos electromagnéticos, las ecuaciones de la difusión de la masa o de la conducción del calor implicaban una velocidad infinita de propagación de los efectos difusivos o caloríficos.
Desde el punto de vista de la física pre-relativista, esto último no planteaba ninguna dificultad, puesto que no se pensaba en que la velocidad de propagación de una perturbación tuviera una cota superior.

Aunque en la actualidad pueda parecer extraño, dada las ideas que tenemos ahora sobre el campo electromagnético, en la época en que Maxwell formuló su teoría tanto él como

la mayor parte de los científicos contemporáneos pensaban que el campo electromagnético se propagaba a través de un medio intangible al que llamaron éter, en forma análoga a como las ondas mecánicas se propagan a través de un sólido elástico.
El éter pasó a cumplir el rol del espacio absoluto en la formulación original de la mecánica. Luego de la crítica de Mach a los fundamentos de la mecánica, se lo identificó con el marco de referencia correspondiente a las estrellas fijas, dotado de propiedades adecuadas para la propagación de ondas electromagnéticas en ausencia de un medio material.
La velocidad de propagación de las ondas electromagnéticas en el vacío, c, que se obtenía a partir de las ecuaciones del campo se interpretaba entonces como la velocidad de propagación respecto del éter. Este último, combinándolo con la idea newtoniana de tiempo absoluto, debería constituir un marco de referencia privilegiado.

De acuerdo al principio de relatividad de Galileo de la mecánica newtoniana, todos los marcos inerciales son equivalentes. Siendo el éter uno de ellos, no cabría esperar que a través de experimentos puramente mecánicos se pudiera detectar un movimiento respecto del éter.
Pero lo que pone aparte al marco inercial ligado al éter de los demás marcos inerciales son, precisamente, los fenómenos electromagnéticos.
Ahora bien, la ecuación de ondas clásica es una consecuencia de las ecuaciones de Maxwell en el vacío, y se verifica para cualquier componente del campo eléctrico o del campo magnético respecto del marco ligado al éter. Entre sus soluciones encontramos ondas planas que se propagan con velocidad c.
Si seleccionamos el eje x en dirección de la propagación de un campo escalar ϕ, la ecuación de ondas para ese campo se reduce a: $\frac{\partial^2 \phi}{\partial x^2} = \frac{1}{c^2} \cdot \frac{\partial^2 \phi}{\partial t^2}$
Si ahora efectuamos un cambio de coordenadas aplicando una transformación de Galileo, la forma de la ecuación de ondas se modifica radicalmente.
El campo $\phi(t,x)$ se expresará ahora como una función $\phi'(t',x')$ de las nuevas coordenadas: $\phi(t,x) = \phi'(t',x')$ Pero, como se desprende de la igualdad precedente, asumimos que sus valores son independientes de las coordenadas de tiempo y espacio: se puede decir que por hipótesis es un invariante respecto de un cambio de coordenadas.
Si la velocidad del nuevo marco de referencia respecto del éter es V, se tiene: $t' = t$
$x' = x - V \cdot t$
En términos de las nuevas coordenadas t' y x' la ecuación de ondas adopta la forma:

$$\left(1 - \frac{V^2}{c^2}\right) \cdot \frac{\partial^2 \phi'}{\partial x'^2} = \frac{1}{c^2} \cdot \frac{\partial^2 \phi'}{\partial t'^2} - \frac{2 \cdot V}{c^2} \cdot \frac{\partial^2 \phi'}{\partial x' \partial t'}$$

Esta ecuación es diferente a la ecuación de ondas clásica. En ella aparece la velocidad relativa del nuevo marco de referencia respecto del marco original. No obstante el cambio de forma, la nueva ecuación posee como soluciones ondas planas en las coordenadas $t' = t$ y x'.
Estas ondas planas se mueven con velocidad $u = c - V$ cuando viajan en el sentido positivo del eje x', y con velocidad $u = c + V$ cuando lo hacen en sentido opuesto al de dicho eje.

Si las transformaciones de Galileo son válidas y si la ecuación de ondas clásica es válida respecto de un marco de referencia privilegiado, el éter, entonces debido a la ley de

adición de las velocidades implícita en las transformaciones de Galileo debería ser posible detectar un movimiento respecto del éste mediante experimentos que involucren la propagación de campos electromagnéticos.
Se concibieron dos tipos de experimentos para intentar una confirmación de la hipótesis del éter.
Unos involucran la propagación de ondas electromagnéticas de frecuencias lumínicas, como los efectuados por Fizeau y por Michelson y Morley.
Otros no involucran dicha propagación, como el experimento con un capacitor cargado y presumiblemente en movimiento respecto del éter, realizado por Trouton y Noble, y los experimentos de Kaufmann de 1902, destinados a la determinación experimental de la inercia aparente de un electrón producido por una desintegración radioactiva de un núcleo atómico (desintegración beta) y que se mueve a una velocidad no despreciable respecto de c (hasta 0.96 c).

En 1851 y 1853 Fizeau midió la pequeña diferencia entre el índice de refracción de la luz en agua en movimiento y en agua en reposo respecto del laboratorio, utilizando un método interferométrico. Sus resultados no coincidieron con lo que se podría haber esperado a partir de la aplicación de la ley de adición de velocidades clásica. No obstante, una aplicación de la ley de adición de velocidades cuando se ve involucrada la propagación de ondas en un medio material es muy discutible[50].
Si v es la velocidad del flujo de agua respecto del laboratorio, en lugar de la velocidad compuesta c + v Fizeau halló la velocidad $c + (1-1/n^2)\ v$, donde n es el índice de refracción del agua. Lorentz logró deducir el coeficiente de convección $(1-1/n^2)$ a partir de la idea de un éter estacionario que no se ve afectado por el movimiento de cargas eléctricas a través de él.

Como es razonable suponer que la Tierra se desplaza respecto del éter[51], cabría esperar la aparición de ciertos efectos ópticos en el vacío debidos en última instancia a diferencias en los tiempos de propagación de la luz entre dos puntos fijos a nuestro planeta, dependiendo de su estado de movimiento en relación con el éter. Así, cuando un rayo de luz parte de un punto A y alcanza un punto B, regresando luego al punto A, el tiempo total insumido tiene que variar cuando ambos puntos se desplazan conjuntamente respecto del éter sin arrastrarlo, en comparación con el tiempo insumido por el mismo proceso cuando los dos puntos permanecen en reposo respecto del éter.
Supongamos que la distancia entre A y B es l. Supongamos que el segmento AB se desplaza con velocidad V respecto del éter, paralelamente a la dirección de la velocidad y con B marchando por delante de A.
Entonces el tiempo que insume el viaje de ida y vuelta de un pulso luminoso de A hacia B y de B hacia A viene dado por $t_1 = \frac{l}{c+V} + \frac{l}{c-V}$, siendo c la velocidad de la luz respecto del éter, y aplicando la ley de adición de velocidades de Galileo a cada uno de los trayectos. Pero si el segmento AB se mueve con esa misma velocidad V perpendicularmente a la dirección de su velocidad, el tiempo insumido en un viaja de

[50] El efecto sería el mismo si la velocidad de movimiento del medio es la misma, para la propagación a través de un medio material cada vez más enrarecido, como el que se puede obtener a partir de un gas.

[51] Suponiendo que la Tierra no lo arrastra durante su movimiento, lo cual parece una hipótesis razonable en vista de las peculiares propiedades que se le asignaban al éter.

ida y vuelta de un pulso luminoso es ahora, aplicando la ley de adición de velocidades y teniendo en cuenta la simetría de los viajes de ida y vuelta: $t_2 = \frac{2 \cdot l}{\sqrt{c^2 - V^2}}$

La diferencia de tiempos es entonces $t_1 - t_2 = \frac{l}{c} \cdot \left(\frac{1}{1+\frac{V}{c}} + \frac{1}{1-\frac{V}{c}} - \frac{2}{\sqrt{1-\frac{V^2}{c^2}}} \right)$

Suponiendo que $\frac{V}{c}$ es pequeña respecto de la unidad, desarrollando en serie de potencias se obtiene la siguiente aproximación: $t_1 - t_2 \cong \frac{l}{c} \cdot \frac{V^2}{c^2}$

Michelson y Morley pusieron en práctica en 1881 una idea que había sido adelantada ya por Maxwell. Midieron la diferencia de fase entre dos haces de luz coherentes, provenientes de la misma fuente, que en una parte de su trayecto hacen un viaje de ida y vuelta por caminos perpendiculares entre sí. Los caminos de luz perpendiculares entre sí que caracterizan el aparato de Michelson y Morley están destinados a producir una diferencia de tiempos que al modificarse produce un corrimiento en las franjas de interferencia que se producen al superponerse los dos haces al final del camino. Para comparar los diferentes casos posibles, el aparato flotaba en un tanque de mercurio y podía rotar muy lentamente. Este es el único tipo de experimento óptico lo bastante sensible como para detectar el efecto debido al éter, que es un efecto de segundo orden: depende como vimos de $\frac{V^2}{c^2}$. Esperaban que al rotar el aparato 90º se produciría un desplazamiento en las franjas de interferencia, excepto si la Tierra se encontrara muy próxima al reposo respecto del éter. Para evitar esta última situación, repitieron el experimento en distintas épocas del año, en las que la velocidad de desplazamiento de nuestro planeta respecto del sol va adoptando direcciones muy diferentes. Repitieron todos los experimentos en 1887, utilizando un aparato modificado para obtener una mayor sensibilidad, obteniendo resultados negativos en ambas instancias.
En 1924 Miller volvió a efectuar el experimento e informó esta vez un efecto positivo (una velocidad de 10 km/s de la Tierra respecto del éter). El tema fue retomado por Joos, trabajando en la planta de la empresa Zeiss en Jena, utilizando un diseño de interferómetro extremadamente sensible. Los experimentos de Joos indican que si hay un efecto positivo, este es menor que una milésima parte del ancho de una franja de interferencia.

El experimento de Trouton y Noble se basa en los estudios de FitzGerald acerca de los fenómenos que se deberían producir cuando un capacitor cargado es arrastrado a través del éter debido al movimiento de la Tierra. Para simplificar, consideremos una carga eléctrica en movimiento respecto del éter. Produce un campo magnético en sus alrededores. Si hay otra carga eléctrica próxima a la primera y que se mueve en forma solidaria con ella, siempre que la velocidad de movimiento de la segunda carga no sea paralela al campo que actúa sobre esa carga, aparecerá una fuerza de origen magnético sobre esta segunda carga. Si las dos cargas, iguales pero de signo opuesto y se encuentran fijas en un sólido, formando un dipolo, este dipolo sufre un momento de torsión que tiende a girar las cargas hasta que el segmento de recta que las une se

dispone perpendicular a la dirección de la velocidad de movimiento respecto del marco de referencia preferencial. Lo mismo cabría esperar cuando las cargas iguales y opuestas se encuentran distribuidas en las placas de un capacitor, el cual tendería a rotar hasta que las placas quedaran paralelas a la velocidad de movimiento respecto del éter.
Si el capacitor se encuentra inicialmente descargado y se desplaza con sus placas paralelas a la dirección del movimiento, FitzGerald demostró que, admitiendo la teoría del éter, en el momento de cargarlo y almacenar así la energía magnética correspondiente al campo asociado a las cargas en movimiento, se debería producir un tirón mecánico sobre el capacitor análogo al que se produciría sobre un cuerpo situado en la superficie de nuestro planeta si su masa aumentara bruscamente de magnitud.
No se pudo medir ni el momento de torsión ni el tirón mecánico, pese a lo cuidadoso de los experimentos de Trouton y Noble.

Por su parte los experimentos de Kaufmann involucraban acciones de campos eléctricos y magnéticos sobre los electrones provenientes de desintegraciones beta, y de ellos resultaba una variación de la masa inercial con la velocidad que difería en cerca de un 20 % respecto de las predicciones de Abraham, obtenidas aplicando la teoría del éter a un modelo de electrón rígido basado en la electrodinámica clásica.

Todos los experimentos que se llevaron a cabo arrojaron resultados negativos: **no hay evidencia experimental que apoye la suposición acerca de la existencia de un marco de referencia preferente para el electromagnetismo.** Como dijo Whittaker en su historia de la electrodinámica clásica[52], **el cuerpo Alfa no parece existir.**
Esta situación desnuda una contradicción entre la mecánica newtoniana y la electrodinámica clásica que la teoría del éter había permitido eludir: o bien la mecánica clásica es correcta y es necesario reformular la teoría electromagnética, o bien esta última es correcta y es necesario revisar la mecánica. Pero las dos teorías no son compatibles.
Intentando compatibilizar la electrodinámica con un marco de referencia absoluto, Lorentz introdujo la hipótesis de un espacio y un tiempo locales dependientes del estado de movimiento de los cuerpos respecto del éter. Los instantes de tiempo y las coordenadas de posición de un evento se conectaban entre sí mediante las así llamadas transformaciones de Lorentz. Los resultados negativos de los experimentos de Michelson y Morley y de Trouton y Noble, así como la variación de la masa con la velocidad en los experimentos de Kaufmann[53] se pueden explicar a partir de las hipótesis efectuadas por Lorentz sin descartar el éter.
Partiendo de un punto de vista muy diferente, Einstein [54] dedujo las transformaciones de Lorentz a partir de un nuevo principio de relatividad que permitió compatibilizar la mecánica con la electrodinámica modificando radicalmente la mecánica newtoniana.
En relación con las relaciones de **compatibilidad** de las teorías físicas, David Hilbert[55]

---

[52] E. Whittaker, "A history of the theories of aether and electricity", volumen II, Thomas Nelson & Sons, Londres, UK, 1953.

[53] Para explicar cuantitativamente los resultados de Kaufmann, Lorentz utilizó un modelo de electrón que se contrae en la dirección del movimiento, en lugar de permanecer rígido como había supuesto Abraham.

[54] También las dedujo Poincaré aplicando un principio de relatividad.

[55] En su trabajo "Axiomatisches Denken" de 1918, (reimpreso en "Gesammelte Abhandlugen", Nº 3, Springer, Berlín, 1935).

llamó la atención sobre lo siguiente: no es suficiente que la estructura de una teoría física, tal como la mecánica, sea internamente consistente, sino que además sus proposiciones no deben contradecir las de las teorías de campos vecinos, como lo es la teoría electromagnética en el caso de la mecánica.
Este requerimiento epistemológico de compatibilidad, también se denominada consistencia externa[56] o coherencia explicativa[57]. Si se admite que las subdivisiones del conocimiento científico son en buena medida convencionales y que hay una unidad subyacente, entonces el requerimiento de compatibilidad debería extenderse a todas las teorías científicas.

5.3 La ecuación de ondas clásica: la tridimensionalidad del espacio físico y la evolución de seres inteligentes

Una consecuencia de las ecuaciones de Maxwell es que las componentes de los campos que se propagan en el vacío deben verificar una ecuación de onda clásica.
Esta ecuación verifica el principio de localidad en su versión fuerte, tal como fue planteado en el capítulo 3 de este trabajo.
Pero resulta que del análisis de las soluciones de la ecuación de onda se desprende que la versión fuerte del principio de localidad encierra a su vez dos variantes, dependiendo del número de dimensiones del espacio físico, y esas variantes parecen tener implicaciones epistemológicas muy diferentes.

Para ver esto, supongamos una propagación en el vacío, a partir de un punto M en el cual, en un instante inicial t_0 saltó una chispa.
En el espacio físico tridimensional la ecuación de ondas clásicas presenta, para un caso como éste, dos soluciones, que corresponden a ondas que se propagan con velocidad c[58]. Una de ellas, denominada solución adelantada, se puede interpretar como una perturbación en forma de cáscara esférica con centro en el punto mencionado y que converge hacia él a medida que transcurre el tiempo.
En ese caso el punto actuaría como un sumidero respecto del campo y no como causa del mismo (que se encontraría en otra parte) por lo cual se debe descartar esta solución.
La otra, denominada solución retardada, corresponde a una perturbación, también en forma de cáscara esférica centrada en el punto donde salta la chispa, pero que diverge a partir de ese punto, punto que respecto de esta solución puede ser considerado como una fuente del campo. Esta solución retardada debe ser retenida sobre la base del principio de causalidad, si la chispa se asume como causa y la cáscara esférica como efecto.

---

[56] M.Bunge, “La investigación científica”, Siglo XXI Eds., Méjico, 2001.

[57] P.Thagard, “The best explanation criteria for theory choice”, Journal of Philosophy, Nº 75, pp 76-92, 1978.

[58] Una discusión sobre las soluciones de la ecuación de ondas clásica en una, dos, tres y en general n dimensiones espaciales puede encontrarse, por ejemplo, en el volumen 2 del libro de R. Courant y D. Hilbert “Methods of mathematical physics”, Interscience, New York, 1962; o en E. Copson, “Partial differential equations”, Cambridge University Press, Cambridge, U. K. , 1975.

Consideremos ahora un punto cualquiera P, siendo d (P, M) la distancia entre el punto donde se produce la chispa y el punto donde se estudia la solución.
La distancia que recorre la perturbación propagada desde el instante t_0 en el que se produce la chispa hasta un instante posterior t es c ($t-t_0$).
Si c ($t-t_0$) < d (P, M) el campo producido por la chispa en el punto P es nulo. También es nulo si c ($t-t_0$) > d (P, M).
Solo cuando c ($t-t_0$) = d (P, M) la respuesta en P es diferente de cero: corresponde al pasaje de la cáscara esférica por ese punto.
Cuando t es anterior a t_0 la respuesta es nula en todo punto del espacio, puesto que la chispa todavía no se produjo.

Debido a la linealidad de la ecuación de ondas clásica, si en lugar de una chispa instantánea, en M se produce una perturbación compleja y duradera, esa perturbación localizada produce la propagación de una onda esférica duradera que en el instante t_0 + (d (P, M) / c) reproduce, atenuado por el efecto de la distancia según el factor 1/ d (p, M), el valor que tenía la perturbación en el punto fuente M en el instante t_0.
Así pues, cualquier historia de una fuente puntual produce un campo a una cierta distancia que reproduce la variación temporal de la fuente: cuando miramos un objeto distante, si nuestros ojos son lo bastante agudos y el efecto de la atmósfera interpuesta es lo bastante tenue, en principio podemos ver exactamente como se está comportando.

Este efecto no se conserva si el espacio físico tuviera solo dos dimensiones, en lugar de tres. En este caso la solución retardada de la ecuación de ondas correspondiente a una chispa instantánea producida en un punto M en el instante t_0 no es una perturbación limitada a una circunferencia centrada en M y de radio c ($t- t_0$), sino que consiste en una perturbación que ocupa también el interior del círculo.
Aparece así un **efecto de cola o efecto de estela** que se propaga acompañando al frente de onda circular desde el lugar donde se produjo la chispa hacia afuera.
Para todo instante t posterior a t_0 y para todo punto P que en el instante t diste de M menos de c ($t-t_0$), el factor de atenuación en este caso es, en el punto P:

$$\frac{2 \cdot c}{\sqrt{c^2 \cdot (t - t_0)^2 - d^2(P, M)}}$$

Si t es anterior a t_0 ese factor es nulo. A medida que t aumenta, el factor de atenuación tiende a cero para un punto P fijo.
Al igual que en un mundo tridimensional, la linealidad de la ecuación de ondas en uno bidimensional tiene como consecuencia que el campo propagado correspondiente a una sucesión de chispas producidas en M en diferentes instantes t_0 es igual a la superposición de los campos que cada chispa produciría por separado.
Pero debido al efecto de estela, una perturbación electromagnética previa a la chispa considerada, es decir producida en un instante t_s **previo** al instante t_0, continuará afectando lo que se observa en el punto P, puesto que el factor de propagación es en este último caso y en el punto P:

$$\frac{2 \cdot c}{\sqrt{c^2 \cdot (t - t_s)^2 - d^2(P, M)}}$$

Si $c \cdot (t - t_0)$ es mayor que d (P, M), entonces $c \cdot (t - t_s) > c \cdot (t - t_0)$ también lo es. Entonces el factor de propagación en el punto P debido a una perturbación en M producida en un instante t_s es distinto de cero para todo $t_s<t_0$.
Como consecuencia de esto, **en un mundo bidimensional, superpuestas a la imagen**

de un objeto lejano aparecerán las imágenes correspondientes a las trazas fantasmas de todos los objetos que alguna vez ocuparon esa misma posición en el pasado. Las ondas producidas por estos objetos continuarán propagándose desde ese punto en forma indefinida.

Lo dicho hasta este momento, si bien concretado para el caso de una chispa y la visión, se aplica a cualquier perturbación que produzca campos que se propaguen de acuerdo con la ecuación de ondas clásica.
La audición, basada en perturbaciones que se propagan de acuerdo a esa ecuación de ondas, también se vería sometida al efecto de cola o de estela en un mundo bidimensional: el sonido reverberaría indefinidamente. Sería como vivir en una cámara de ecos.
Es posible demostrar que el efecto de estela aparecería también en un espacio de cuatro dimensiones[59].
Todo esto sugiere que los mundos bidimensionales o tetra dimensionales no serían favorables para la evolución de seres inteligentes, posiblemente debido a las dificultades en la percepción y en la comunicación, generadas por la superposición de información en el receptor asociada con las trazas fantasmas.

[59] Estas dimensiones son todas espaciales, el tiempo se halla excluido. Una discusión de estos efectos y su relación con el principio de Huygens de la óptica clásica, puede hallarse en el artículo de M. Gutzwiller "Huygens' principle and the path integral" en S. Lundquist y otros (Eds.) "*Path summation: achievements and goals*", World Scientific, Singapore, 1988, pp. 47-73. Las consecuencias de la tridimensionalidad del espacio físico ha sido y continúa siendo muy investigada en el marco de la física teórica. Un ejemplo clásico de estas investigaciones es el trabajo de Paul Ehrenfest "¿Cómo las leyes fundamentales de la física ponen en evidencia que el espacio posee tres dimensiones?", publicado en alemán en los Annalen der Physik, vol. 366, Nº 5, 1920, pp. 440-446. Si la dimensión del espacio es mayor que tres y se generalizan a ese espacio las leyes de interacción gravitatoria y electroestática, entonces los resultados principales obtenidos hasta este momento se pueden resumir así: (1) La mecánica clásica permite demostrar que las órbitas cerradas, todavía posibles, ya no son estables (cualquier perturbación las destruye) de modo que los sistemas planetarios no pueden formarse y mantenerse tal como los conocemos. Este resultado astronómico continúa verificándose cuando se analiza la estabilidad de las órbitas en el marco de la teoría generalizada de la relatividad. (2) La mecánica cuántica permite concluir que los átomos y moléculas, considerados como configuraciones estables de electrones y núcleos, no pueden existir. Admitiendo estos resultados, podemos concluir que la dimensión del espacio físico debe ser inferior a cuatro. El argumento sobre las trazas fantasmas sugiere que los espacios de una y dos dimensiones también podrían excluirse, con lo cual nos quedaríamos con el espacio tridimensional.

6. Los conceptos de tiempo y espacio en la teoría restringida de la relatividad

"*Fue gracias al electromagnetismo que se creó la relatividad restringida. Los sabios de aquella época querían hallar una transformación entre el espacio y el tiempo que permitiera mantener invariantes las ecuaciones de Maxwell del electromagnetismo al pasar de un sistema de referencia a otro en traslación uniforme respecto del primero. Es Poincaré quien, apoyándose en el electromagnetismo, obtiene la transformación correcta, llamada de ahí en adelante transformación de Lorentz-Poincaré. Esta transformación permite obtener la invariancia deseada, como lo muestra Poincaré en su famosa publicación del 5 de junio de 1905….Este origen electromagnético de la relatividad restringida, retomada por Einstein, fue rápidamente criticado. Los físicos no admitían que un fenómeno físico particular pudiera engendrar una super-ley entre el tiempo y el espacio -la transformación de Lorentz-Poincaré- aplicable a todas las leyes de la física. Era necesario, pues, buscar ideas de bases más generales para demostrar la famosa transformación, ideas referidas esencialmente al espacio y el tiempo. En 1910, un primer trabajo de investigación en esa dirección fue realizado por Waldemar Ignatowski. En el curso del siglo XX, esas nuevas ideas desembocaron en una demostración de la transformación de Lorentz-Poincaré basada sobre hipótesis independientes de todo fenómeno electromagnético.*"

J. Hladik « Pour comprendre simplement les origines et l'évolution de la physique quantique », capítulo 9, Ellipses, Paris, 2008.

En 1904 Henri Poincaré, en una famosa conferencia sobre los principios de la física-matemática[60], analizó el estado de cosas en ese entonces, señaló las dificultades y planteó la posibilidad de construir una nueva mecánica en sustitución de la mecánica clásica. Revisó los numerosos e infructuosos esfuerzos por salvar el paradigma establecido, añadiendo hipótesis tras hipótesis, como la del tiempo local de Lorentz y la contracción de la dimensión de un cuerpo paralelamente a su dirección de movimiento (contracción de Fitz-Gerald). En esa misma conferencia, continuando sus reflexiones sobre la medición del tiempo iniciadas en 1898[61], Poincaré plantea un método para sincronizar relojes situados en puntos diferentes de un mismo marco de referencia. Esta misma idea es retomada por Einstein, nueve meses después, en su famoso artículo de 1905 sobre la electrodinámica de los cuerpos en movimiento que se considera el trabajo fundador de la teoría restringida de la relatividad.

En un trabajo publicado durante 1904 [62], Poincaré enunció un principio de relatividad estableciendo que **todas** las leyes físicas, y en particular las ecuaciones matemáticas que describen estas leyes, toman la misma forma general en todos los sistemas de referencia que se encuentren en movimiento relativo rectilíneo y uniforme.

En el mencionado artículo de 1905 Einstein enuncia un principio de relatividad

[60] H. Poincaré, "Los principios de la física matemática", en Albert Einstein y otros, "*La teoría de la relatividad*", Altaya, Barcelona, 1993.

[61] H. Poincaré, ´´La mesure du temps´´, Revue de métaphysique et de morale, 6, pp. 371-384, 1898.

[62] H. Poincaré, ´´L'état actuel et l'avenir de la physique mathématique´´, Bulletin des Sciences Mathématiques, (2) 28, pp. 302-324,1904. En los libros de J. Hladik "*Pour comprendre simplement la théorie de la relativité*", Ellipses, París, 2005 y de A.Logunov "Henri Poincaré and relativity theory", Moscú, Nauka, 2005 se puede hallar, respectivamente, un estudio histórico muy completo y una presentación teórica del desarrollo de la teoría restringida y el rol de iniciador que le cupo a Poincaré.

esencialmente equivalente al propuesto por Poincaré, complementándolo con un postulado sobre la invariancia de la velocidad de la luz.

El principio se puede resumir en estos dos enunciados:
(a) Las leyes de la electrodinámica (incluyendo un valor constante c para la propagación del campo electromagnético en vacío) y de la mecánica clásica son las mismas en todos los marcos de referencia inerciales[63].
(b) No es posible proyectar un experimento que permita definir un estado de movimiento absoluto en el sentido newtoniano.

A partir del principio de relatividad, la contradicción entre la mecánica y la electrodinámica fue resuelta finamente en 1905, dando prioridad a la electrodinámica sobre la mecánica y planteando las bases de una nueva mecánica fundada en una revisión radical de los conceptos físicos de tiempo y espacio que se habían aceptado hasta ese momento.

6.1 El principio de relatividad de Einstein, las transformaciones de Lorentz y sus consecuencias cinemáticas

Asumiendo que el campo electromagnético no necesita de un medio para propagarse, que es una realidad en sí, se comprende que las ecuaciones de la electrodinámica clásica puedan ser las mismas respecto de cualquier marco de referencia inercial.
En ese caso la velocidad c debe ser la misma para cualquier observador, independiente de su estado de movimiento respecto de la fuente del campo: este es el postulado de constancia de la velocidad de la luz introducido por Einstein y parece confirmado por los resultados de los experimentos como el de Michelson-Morley.
El principio de relatividad se expresa a través del postulado siguiente: las leyes físicas se deben formular de modo que adopten la misma forma matemática (sean igualmente válidas) en todos los marcos inerciales.
El principio de la constancia de la velocidad de la luz podría resultar compatible con el principio de relatividad si éste no se expresara a través de las transformaciones de Galileo: estas últimas implican una ley de adición de velocidades inaplicable en electrodinámica.
Para evitar confusiones, resulta conveniente denominar **principio de relatividad de Einstein** a la combinación del principio de relatividad con postulado de una velocidad máxima de propagación de las interacciones. La mecánica compatible con este principio se denomina **mecánica relativista**.

Reflexionando sobre cómo se mide el tiempo en Física, Einstein se dio cuenta de ciertas hipótesis no justificadas por la experiencia que se encuentran en la base del concepto de tiempo absoluto de la mecánica clásica[64].
La asignación de instantes de tiempo a diferentes eventos que ocurren en un mismo lugar respecto de un marco de referencia no presenta problema, al igual que el concepto de simultaneidad de dos acontecimientos que se producen en el mismo punto. El

[63] El sistema utilizado en astronomía se considera un marco inercial. Los demás se muevan con movimiento rectilíneo y uniforme respecto de este sistema, y por tanto los unos respecto de los otros.

[64] A. Einstein "*La Relatividad. Memorias originales*" Emecé, Buenos Aires, 1950.

problema aparece cuando se intenta determinar si dos sucesos que se producen en lugares alejados son o no simultáneos, porque para hacer esto es necesario sincronizar los relojes situados en esos lugares.
Parece de sentido común suponer que para sincronizar un reloj en A con un reloj alejado en B basta con sincronizar primero un reloj U_B situado en A mediante el reloj U_A asignado al punto A, y luego transportar el reloj U_B desde A hasta B, dejándolo en B y leyendo desde ese momento en adelante el tiempo de B en U_B.
Esta forma de sincronizar relojes resulta aplicable si ocurre lo siguiente: siempre que dos relojes U_{B1} y U_{B2}, sincronizados en A, se pueden transportar hasta B por caminos diferentes pero arribando simultáneamente a destino, entonces resultan estar sincronizados en B. Si se admite a priori, puede denominarse **postulado del transporte.**
Si la sincronización no se conserva al llegar a B, cuando los relojes son transportados a partir de A por caminos diferentes, entonces es necesario introducir un método nuevo para sincronizar relojes situados en puntos distantes.
El concepto de simultaneidad de eventos que aparece en el artículo de Einstein sobre le electrodinámica de los cuerpos en movimiento, se basa en admitir que el método del transporte no es válido[65] y en sustituirlo por un método basado en el envío de señales entre los puntos cuyos relojes se desea sincronizar.
La posibilidad de este segundo método reposa a su vez en una suposición adicional: ninguna cadena causal que comienza en A puede dejar sentir sus efectos en un punto B **antes** de la llegada de una señal luminosa (onda electromagnética) que es emitida simultáneamente con el inicio de la cadena en A y viaja en el vacío hasta el punto B. Aquí "antes" se refiere a un orden cronológico determinado con un reloj en B, que para esta tarea no necesita estar sincronizado con el reloj de A.
En principio se podrían conectar puntos distantes a través de señales de distinta naturaleza, incluso en el vacío, pero este postulado de existencia de un tipo de señal de máxima velocidad de propagación, a través de un enunciado que involucra toda conexión causal posible, singulariza las señales luminosas y las pone aparte de las otras señales posibles.
Puesto que la velocidad de la luz parece ser independiente del movimiento de la fuente y la misma para todos los observadores situados en marcos inerciales, Einstein propone un método de sincronización basado en el envío y recepción de señales luminosas.
Consideremos dos relojes situados en dos puntos A y B fijos respecto de un marco de referencia K. Supongamos que en el instante $t_{A,1}$ una señal luminosa parte de A hacia B, se refleja allí y retorna con información sobre el instante t_B en el que se produjo la reflexión en B según el reloj allí situado. La señal reflejada arriba al punto A cuando el reloj allí indica el instante $t_{A,2}$. Por definición, los relojes están sincronizados cuando

$$t_B = \frac{1}{2} \cdot \left(t_{A,1} + t_{A,2} \right)$$

Dos eventos que ocurren en puntos diferentes de un mismo marco de referencia son simultáneos cuando los relojes locales en los lugares correspondientes marcan el mismo instante de tiempo.

Esta forma de definir sincronismo de relojes se puede re-escribir así: $t_B - t_{A,1} = t_{A,2} - t_B$
Pero, como señala Einstein en "Sobre la electrodinámica de los cuerpos en

[65] Suponía que, en general, las diferencias en las trayectorias elimina el sincronismo de los relojes en el punto de arribo.

movimiento", esto implica la suposición que el "tiempo" que requiere la luz para viajar de A hasta B es igual al "tiempo" que requiere para viajar de B hasta A[66].
Además, es necesario suponer que esta definición de sincronismo se puede aplicar para todos los puntos en reposo en un mismo marco de referencia, sin contradicciones: la relación de sincronismo debe asumirse simétrica (si A está sincronizado con B entonces B está sincronizado con A) y transitiva (si A está sincronizado con B y B lo está con C, entonces A está sincronizado con C).

Supongamos ahora que en el punto medio C del segmento que une los puntos A y B se ha colocado una fuente de luz con su correspondiente reloj sincronizado. En un instante dado desde C se emite un destello luminoso que, de acuerdo con lo establecido previamente, alcanza los puntos A y B en forma simultánea.
Respecto de un observador en un marco de referencia K' respecto del cual K se encuentra en movimiento paralelamente al segmento AB, las señales luminosas se propagan con la misma velocidad, pero mientras viajan a partir de C, uno de los puntos (A o B) se aleja de la señal que viaja hacia él mientras que el otro punto (B o A) se acerca.
Como consecuencia, pese a la simetría geométrica, una de las señales alcanza uno de los extremos del segmento antes que la otra pueda alcanzar el suyo.
Así pues, la simultaneidad de dos sucesos que se producen en lugares diferentes es un concepto relativo al observador: dos sucesos que son simultáneos para un observador (K) pueden no serlo para otro observador (K').
Este es el resultado fundamental en relación con el tiempo que se desprende del postulado de la constancia de la velocidad de la luz.
Pero si la simultaneidad de dos eventos es relativa al marco de referencia, entonces la ecuación $t' = t$ que iguala los instantes en los que se produce un mismo evento en dos marcos de referencia tiene que ser falsa[67]. Como la ley de adición de velocidades de Galileo también es falsa, es preciso reexaminar la totalidad de la conexión entre las coordenadas de tiempo y espacio de un mismo evento respecto de dos marcos de referencia inerciales.

Las ecuaciones para la transformación del tiempo y las coordenadas espaciales desde el marco K al marco K':

(1) Deben mantener invariante su forma al ser invertidas, es decir al regresar del marco K' al marco K invirtiendo el signo de la velocidad relativa V.
(2) Deben convertir regiones acotadas de puntos (t, x, y, z) de K en regiones acotadas de puntos (t', x', y', z') de K'.
(3) Cuando V tiende a cero la transformación debe tender a la identidad.

(4) La ley de adición de velocidades resultante de esa transformación debe dejar invariante la velocidad de la luz en el vacío: $c = c'$

[66] A. Einstein "La Relatividad. Memorias originales" Emecé, Buenos Aires, 1950.

[67] Puesto que de ella se desprende que la simultaneidad de dos eventos es independiente del marco de referencia, es decir, del observador.

Es posible demostrar que las restricciones (1) y (2) tomadas en conjunto implican que la transformación deba ser lineal. De la linealidad de la transformación y de (1), (3) y (4) se obtienen las transformaciones de Lorentz. Cuando los marcos inerciales poseen un eje de coordenadas paralelo a la dirección de su movimiento relativo **y los orígenes de tiempo y espacio coinciden en un mismo evento** (Figura 1), se deducen las expresiones más simples de la transformación de Lorentz:

$$x' = \frac{x - V.t}{\sqrt{1 - \frac{V^2}{c^2}}} \qquad y' = y \qquad z' = z \qquad t' = \frac{t - \frac{V.x}{c^2}}{\sqrt{1 - \frac{V^2}{c^2}}}$$

Si $\frac{V}{c}$ es despreciable respecto de 1, las fórmulas se reducen transformaciones de Galileo: $x' = x - V.t \qquad y' = y \qquad z' = z \qquad t' = t$

En el límite de las velocidades bajas respecto de la velocidad de la luz, para todos los fines prácticos el tiempo se comporta como un invariante al pasar de un marco de referencia a otro.

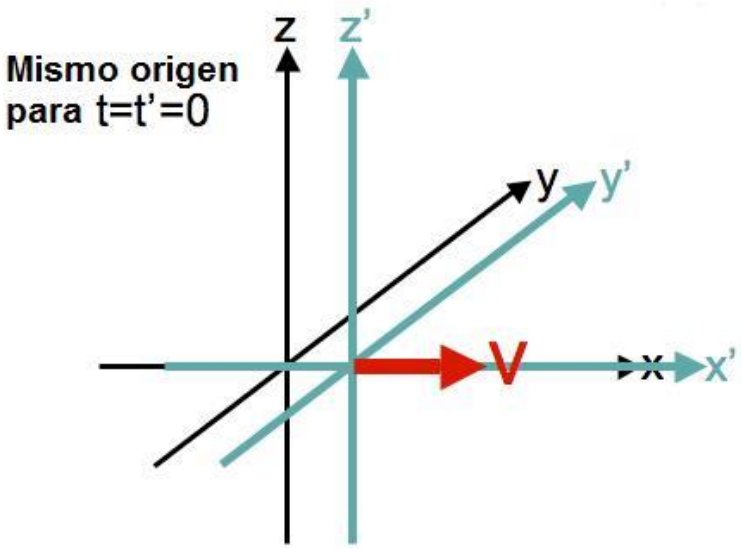


Figura 6.1. El marco K' se desplaza con velocidad V respecto del marco K

Invirtiendo las fórmulas que conectan K con K' se obtiene:

$$x = \frac{x' + V.t'}{\sqrt{1 - \frac{V^2}{c^2}}} \qquad y = y' \qquad z = z' \qquad t = \frac{t' + \frac{V.x'}{c^2}}{\sqrt{1 - \frac{V^2}{c^2}}}$$

Si un reloj se encuentra fijo al origen de K', entonces $x' = 0$ para todo t'.

De la última fórmula se desprende que $t = \frac{t'}{\sqrt{1 - \frac{V^2}{c^2}}}$ o bien $t' = \sqrt{1 - \frac{V^2}{c^2}} \cdot t < t$

Esto significa que un reloj en movimiento respecto de un observador fijo en K, atrasa

siempre respecto a un reloj equivalente en reposo respecto del observador[68].
Obsérvese que el tiempo t es el que indica un reloj sincronizado en K que en ese instante se encuentra a una distancia $x = V \cdot t$ del origen de K.

La llegada de los mesones pi al suelo constituye un ejemplo del atraso de un reloj en movimiento. Esos mesones se producen en reacciones nucleares cuando los rayos cósmicos provenientes del espacio exterior chocan con los núcleos de los átomos e iones atmosféricos a unos 20 km de altura. En el laboratorio los mesones pi poseen un tiempo de vida media de $2{,}55\times10^{-8}$ s, por lo cual, si viajaran a la velocidad de la luz se desintegrarían la mitad de los mesones pi remanentes cada 7,6 m de recorrido. Entonces, luego de 20 km deberían de haber transcurrido $(2\times10^4/0{,}8) = 2\,500$ vidas medias.
La fracción que alcanzaría el suelo sería entonces $(1/2)^{2.500} \approx 10^{-750}$. Este número es extraordinariamente pequeño. Se deberían producir 10^{750} mesones para que uno pudiera llegar al suelo. El número de átomos del universo podría ser 10^{80}, así que no alcanza la totalidad de sus núcleos para asegurar que un mesón pi llegue al suelo.
¿Qué pasa entonces? Los mesones pueden ser considerados como relojes en movimiento que se desplazan a velocidades próximas a c. Para un observador en el suelo el reloj del mesón atrasa con un factor $\sqrt{1-\frac{V^2}{c^2}}$ tan pequeño que muchos de ellos alcanzan el suelo antes de desintegrarse. Sin este atraso relativista un haz de mesones pi se desvanecería luego de 100 m de recorrido. El empleo de grandes aceleradores de partículas permite una confirmación experimental del atraso de un reloj en movimiento. Con un sincrotrón se pueden producir mesones pi con velocidades lo bastante próximas a c como para obtener haces de piones que alcanzan varios centenares de metros a partir del lugar en el que se forman.

Supongamos que una varilla está en reposo respecto de K', situada a lo largo del eje x', y que un observador en K' mide su longitud obteniendo $x_2' - x_1' = \Delta x'$ Ahora, K' se mueve respecto de K. Si un observador en K pretende medir la longitud de la varilla sin detenerla, debe determinar con qué puntos x_1 y x_2 del eje x coinciden los extremos de la varilla en un mismo instante t. Entonces considera dos sucesos que respecto de K son simultáneos. Como de las transformaciones de Lorentz se desprende que

$$x_2' = \frac{x_2 - V.t}{\sqrt{1-\frac{V^2}{c^2}}} \qquad x_1' = \frac{x_1 - V.t}{\sqrt{1-\frac{V^2}{c^2}}}$$

Restando miembro a miembro se obtiene:

$$x_2' - x_1' = \Delta x' = \frac{\Delta x}{\sqrt{1-\frac{V^2}{c^2}}} \quad \text{o sea} \quad \Delta x = \sqrt{1-\frac{V^2}{c^2}} \cdot \Delta x'$$

Una varilla en movimiento resulta más corta que en reposo: los cuerpos en movimiento se acortan. Obsérvese que si bien los sucesos (t, x_1) y (t, x_2) son simultáneos respecto del marco K, no lo son respecto de K' en el cual la varilla se encuentra en reposo.

Retomando el ejemplo de los mesones pi, un observador fijo a los mesones mediría el mismo tiempo de vida media que se determina en el laboratorio. Pero como se mueve a

[68] Un reloj en reposo respecto de un marco de referencia se denomina el reloj del observador para ese marco.

una velocidad V respecto de la atmósfera, los $\Delta x' = 20km$ de espesor de esta última para él son $\Delta x = \sqrt{1-\frac{V^2}{c^2}} \cdot \Delta x'$. Como V se encuentra próxima a c, esta distancia es tan corta que permite que el mimo número de mesones pi que estima el observador fijo a la tierra alcancen el suelo. El resultado es el mismo para ambos observadores.

La dilatación del tiempo y el acortamiento de las longitudes en la dirección del movimiento son efectos puramente cinemáticos, debidos a la peculiar forma en la que se interrelacionan tiempo y espacio según el principio de relatividad de Einstein: en todo esto las interacciones entre los cuerpos no tienen relevancia alguna.

A partir de las transformaciones de Lorentz se obtienen las siguientes expresiones para la adición de velocidades:

$$v'_x = \frac{v_x - V}{1-\frac{V \cdot v_x}{c^2}} \qquad v'_y = \frac{v_y \cdot \sqrt{1-\frac{V^2}{c^2}}}{1-\frac{V \cdot v_x}{c^2}} \qquad v'_z = \frac{v_z \cdot \sqrt{1-\frac{V^2}{c^2}}}{1-\frac{V \cdot v_x}{c^2}}$$

$$v_x = \frac{v'_x + V}{1+\frac{V \cdot v'_x}{c^2}} \qquad v_y = \frac{v'_y \cdot \sqrt{1-\frac{V^2}{c^2}}}{1+\frac{V \cdot v_x}{c^2}} \qquad v_z = \frac{v'_z \cdot \sqrt{1-\frac{V^2}{c^2}}}{1+\frac{V \cdot v_x}{c^2}}$$

Si $v = \sqrt{v_x^2 + v_y^2 + v_z^2} = c$ entonces $v' = \sqrt{v_x'^2 + v'^2_y + v'^2_z} = c$ como cabía esperar.

Esto permite explicar el resultado negativo del experimento de Michelson y Morley.

Si V es despreciable respecto de c se obtienen las fórmulas de adición de velocidades de la mecánica clásica. Además, la combinación de velocidades inferiores a c siempre da como resultado una velocidad inferior a c.

La discrepancia entre los resultados experimentales y la teoría del éter en el caso de los experimentos de Fizeau se explica también mediante la nueva ley de adición de velocidades. Supongamos un medio con índice de refracción n, en este caso agua.

Si la luz en el agua viaja con velocidad $v'_x = \frac{c}{n}$ $\quad v'_y = 0$ $\quad v'_z = 0$ mientras que el agua se mueve paralelamente al eje x con velocidad V, entonces las componentes de la velocidad de la luz respecto del laboratorio vienen dadas por las ecuaciones:

$$v_x = \frac{\frac{c}{n}+V}{1+\frac{V}{c \cdot n}} \approx \left(\frac{c}{n}+V\right) \cdot \left(1-\frac{V}{c \cdot n}\right) \approx \frac{c}{n} + V \cdot \left(1-\frac{1}{n^2}\right) \qquad v_y = 0 \qquad v_z = 0$$

Se obtiene como aproximación la expresión hallada experimentalmente por Fizeau, incluyendo su coeficiente de convección.

Obsérvese que lo que en estos ejemplos se adiciona es una velocidad de movimiento de un cuerpo respecto de un marco de referencia K, con una velocidad -V de movimiento del marco de referencia K respecto de otro marco K' para calcular la velocidad de

movimiento de ese mismo cuerpo respecto de K'. Dicho de otra forma, la velocidad de un cuerpo A (que en este caso coincide con el marco K) referida a un marco K' se combina con la velocidad de un cuerpo B respecto de A (es decir, K) para obtener la velocidad de B respecto de K'.
Si lo que tenemos es la velocidad de **un cuerpo A** respecto de un marco K y la velocidad de **otro cuerpo B** respecto del **mismo marco** K, entonces **la velocidad relativa del cuerpo A respecto del cuerpo B** se calcula según la fórmula clásica, porque todas las operaciones se realizan respecto del mismo marco de referencia.
Si bien la velocidad c es un límite superior para todas las velocidades que pueden ser alcanzadas por los cuerpos materiales [69], la velocidad relativa de un cuerpo respecto de otro en el mismo marco de referencia puede resultar mayor que c. Lo que ocurre es que de esto no se desprende que exista un cuerpo que pueda moverse respecto de K con esa velocidad mayor que c: no solamente en este caso no lo hay, sino que según la teoría de la relatividad no puede haberlo.

6.2 Homogeneidad del tiempo, homogeneidad e isotropía del espacio y principio de relatividad de Einstein

Por razones que se verán posteriormente, resulta conveniente asociar a cada par de eventos, respecto de cada marco de referencia K, una cantidad escalar Δs denominada intervalo entre eventos y definida por la fórmula:

$$\Delta s^2 = c^2 \cdot \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2$$

Aquí $\Delta t = t_2 - t_1$, $\Delta x = x_2 - x_1$, $\Delta y = y_2 - y_1$ y $\Delta z = z_2 - z_1$ son las diferencias de coordenadas temporales y espaciales entre los eventos considerados. El intervalo conecta las coordenadas de tiempo y espacio, que entonces se consideran en conjunto. Esto condujo a Hermann Minkowski a introducir un espacio abstracto de cuatro dimensiones denominado espacio-tiempo, que se estudia en el capítulo próximo.
Al mismo par de eventos, considerados ahora respecto de otro marco K', les corresponderá un intervalo construido en base a las correspondientes coordenadas de esos eventos:

$$\Delta s'^2 = c^2 \cdot \Delta t'^2 - \Delta x'^2 - \Delta y'^2 - \Delta z'^2$$

A partir de las transformaciones de Lorentz se desprende que $\Delta s^2 = \Delta s'^2$ Es decir, el intervalo entre dos eventos es independiente del marco de referencia inercial considerado.

Sobre la base de los postulados de homogeneidad del espacio y el tiempo, y de isotropía del espacio, es posible demostrar la invariancia del intervalo entre dos eventos en forma directa, sin utilizar las transformaciones de Lorentz, para luego deducir estas últimas a partir de la invariancia del intervalo. A continuación revisaremos esa deducción.
El contenido de esta sección podría resultar más demandante, para algunos lectores, que el resto del capítulo. Si lo desea, el lector puede omitirlo y pasar directamente a la sección 6.3.

Del postulado de la constancia de la velocidad de la luz se desprende que si el intervalo

[69] Las fórmulas de Lorentz ya sugieren esto, desde que $\sqrt{1-\frac{V^2}{c^2}}$ se haría imaginario. Como se desprende del análisis de la sección 8.1, la dinámica relativista conduce a la misma conclusión.

se anula respecto de K, se anula respecto de K', puesto que los eventos se encuentran conectados por una señal que viaja con velocidad c.

Si $ds = \sqrt{c^2 \cdot dt^2 - dx^2 - dy^2 - dz^2}$ es el intervalo entre dos eventos respecto de K, eventos que se toman arbitrariamente próximos en el tiempo y en el espacio, y $ds' = \sqrt{c^2 \cdot dt'^2 - dx'^2 - dy'^2 - dz'^2}$ es el correspondiente intervalo respecto de K', entonces $ds = 0$ implica $ds' = 0$.

Como ds y ds' deben ser infinitésimos del mismo orden y se anulan a la vez, cabe esperar que sean proporcionales: $ds = \alpha \cdot ds'$

El coeficiente de proporcionalidad no puede depender de las coordenadas de tiempo y espacio. Si dependiera de ellas, diferentes puntos del espacio o diferentes instantes de tiempo no serían equivalentes, lo cual violaría el postulado de la homogeneidad del espacio y del tiempo.

Tampoco puede depender de la dirección ni del sentido de la velocidad relativa de K' respecto de K, puesto que si así fuera se violaría el postulado de la isotropía del espacio.

Entonces: $\alpha = \alpha\left(\left|\vec{V}\right|\right)$ Pero: $\left|\vec{V}\right| = \left|-\vec{V}\right|$ Así pues, resulta que no solo se verifica $ds = \alpha \cdot ds'$ sino que también se verifica: $ds' = \alpha \cdot ds$.

Como consecuencia: $\alpha^2 = 1$ Es decir, $\alpha = \pm 1$

Como se puede suponer que $\alpha = \alpha\left(\left|\vec{V}\right|\right)$ es función continua de $\left|\vec{V}\right|$, o bien $\alpha = 1$ o bien $\alpha = -1$ en todos los casos. Pues si para un valor V_1 de $\left|\vec{V}\right|$ fuera $\alpha = 1$ y para otro valor V_2 fuera $\alpha = -1$, para valores comprendidos entre V_1 y V_2, $\alpha = \alpha\left(\left|\vec{V}\right|\right)$ debería tomar todos los valores entre -1 y +1.

Por otra parte, cuando K'=K, debe ser $ds' = ds$, de modo que si $\alpha = \alpha\left(\left|\vec{V}\right|\right)$ ha de ser constante, su valor debe ser necesariamente 1.

Por tanto, si $\alpha = \alpha\left(\left|\vec{V}\right|\right) = 1$, para todo par de marcos inerciales se verifica la igualdad de los intervalos $ds' = ds$.

De esta última igualdad se desprende que en general $\Delta s' = \Delta s$

Así pues el intervalo entre eventos es un invariante: es el mismo respecto de todos los marcos de referencia inerciales.

A partir de la invariancia del intervalo se pueden deducir las fórmulas de Lorentz.

Para hacer esta deducción se trabaja en forma geométrica en el tiempo-espacio de Minkowski. Las transformaciones de coordenadas que dejan invariante el intervalo y conectan dos marcos inerciales K y K' entre sí, cuyos orígenes de coordenadas coinciden, se pueden interpretar como rotaciones en el espacio de cuatro dimensiones. Los coeficientes de la correspondiente matriz de cambio de coordenadas se pueden calcular en términos de la velocidad de un marco respecto del otro y de la velocidad de la luz c[70].

Para deducir las transformaciones de Lorentz conviene introducir una coordenada imaginaria $\tau = i \cdot c \cdot t$ con lo cual el cuadrado del intervalo entre dos eventos, referidos a

[70] Se puede encontrar una deducción de las transformaciones de Lorentz siguiendo el camino aquí esbozado en L. Landau y E. Lifchitz, "*Teoría del Campo*", Reverté, Barcelona, 1982.

un marco K verifica $-\Delta s^2 = \Delta\tau^2 + \Delta x^2 + \Delta y^2 + \Delta z^2$, y respecto de otro marco K' se tiene $-\Delta s'^2 = \Delta\tau'^2 + \Delta x'^2 + \Delta y'^2 + \Delta z'^2$

Debido a que el intervalo permanece invariante, se verifica para dos eventos dados y sus coordenadas respecto de dos marcos inerciales:

$$\Delta\tau^2 + \Delta x^2 + \Delta y^2 + \Delta z^2 = \Delta\tau'^2 + \Delta x'^2 + \Delta y'^2 + \Delta z'^2$$

En general se puede suponer que las coordenadas de un evento respecto de K' son funciones regulares y biunívocas de las coordenadas de ese mismo evento respecto de K:

$$\tau' = \tau'(\tau, x, y, z) \qquad x' = x'(\tau, x, y, z) \qquad y' = y'(\tau, x, y, z) \qquad z' = z'(\tau, x, y, z)$$

$$\tau = \tau(\tau', x', y', z') \qquad x = x(\tau', x', y', z') \qquad y = y(\tau', x', y', z') \qquad z = z(\tau', x', y', z')$$

Estas son las fórmulas de transformación de coordenadas en el espacio tetra-dimensional, y equivalen a las transformaciones de Lorentz, con la diferencia que implica haber introducido la coordenada temporal imaginaria $\tau = i \cdot c \cdot t$

Partiendo de la igualdad diferencial $d\tau^2 + dx^2 + dy^2 + dz^2 = d\tau'^2 + dx'^2 + dy'^2 + dz'^2$ es posible demostrar que todas las derivadas parciales de segundo orden (incluyendo todas las derivadas cruzadas) de esas funciones, respecto de las variables (τ, x, y, z), son nulas. Como consecuencia, se puede demostrar que las coordenadas respecto de cualquier evento respecto de K' se pueden expresar como la suma de una traslación más una rotación en términos de las coordenadas de ese mismo evento respecto de K [71].

La traslación no presenta interés puesto que simplemente ubica respecto de K' el origen de coordenadas de K, de modo que se puede ignorar suponiendo que ambos orígenes de coordenadas coinciden. Queda así la rotación en cuatro dimensiones, que se puede descomponer en seis rotaciones en los planos **xy**, **yz**, **zx**, **τx**, **τy**, y **τz**, en forma análoga a como una rotación en el espacio tridimensional se puede resolver en tres rotaciones en los planos **xy**, **yz**, **zx.**

Las rotaciones en los planos **xy**, **yz**, **zx** transforman solamente las coordenadas espaciales sin involucrar el tiempo, por lo cual corresponden a las rotaciones ordinarias en el espacio físico y no presentan interés. Lo más sencillo es analizar una rotación en uno de los planos **τx**, **τy**, o **τz.**

Consideremos entonces una rotación en el plano **τx**, suponiendo que las coordenadas y,z permanecen incambiadas: esta situación corresponde al caso en el que el marco K' se traslada respecto de K con velocidad uniforme V de modo que x' y x son paralelos.

Si φ es el ángulo de rotación, entonces:

$$x = x' \cdot \cos\varphi - \tau' \cdot sen\varphi \qquad\qquad \tau = x' \cdot sen\varphi + \tau' \cdot \cos\varphi$$

Si el origen de K' se mueve con velocidad V respecto de K, poniendo x'=0 en las fórmulas de rotación se obtiene $x = -\tau' \cdot sen\varphi \qquad \tau = \tau' \cdot \cos\varphi$ de donde $\frac{x}{\tau} = -tg\varphi$

Como $\tau = i \cdot c \cdot t$ y en este caso $\frac{x}{t} = V$, resulta que $tg\varphi = i \cdot \frac{V}{c}$

De esta última fórmula y de relaciones trigonométricas bien conocidas se desprende que

[71] Ver J.Synge "*Relativity: the special theory*", North-Holland, Amsterdam, 2ª edición, 1965, pp. 69-70.

$$sen\varphi = \frac{\frac{iV}{c}}{\sqrt{1-\frac{V^2}{c^2}}} \quad \text{y} \quad \cos\varphi = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}}$$

Sustituyendo en las fórmulas para la rotación $x = \dfrac{x' - i\frac{V}{c}\tau'}{\sqrt{1-\frac{V^2}{c^2}}}$ $\qquad \tau = \dfrac{x' + i\frac{V}{c}x'}{\sqrt{1-\frac{V^2}{c^2}}}$

Regresando al tiempo real: $x = \dfrac{x' + V \cdot t'}{\sqrt{1-\frac{V^2}{c^2}}}$ $\qquad t = \dfrac{t' + \frac{V}{c^2}x'}{\sqrt{1-\frac{V^2}{c^2}}}$ $\qquad y = y'$ $\qquad z = z'$

Estas son las fórmulas de Lorentz.

6.3 Formulación de la teoría de la relatividad independiente del postulado de constancia de la velocidad de la luz: la constante de estructura del tiempo-espacio

Desde el punto de vista histórico la relatividad restringida surgió en íntima conexión con la teoría electromagnética. La deducción de las transformaciones de Lorentz vista en 6.2 se apoya en la invariancia de la velocidad de propagación de las ondas electromagnéticas al pasar de un marco inercial a otro.
No obstante, las transformaciones en sí mismas conectan las coordenadas de tiempo y espacio de un mismo evento respecto a marcos de referencia inerciales: según el principio de relatividad son aplicables a todas las leyes físicas, más allá de su origen electromagnético.

Por este motivo, durante buena parte del siglo pasado se realizaron numerosos esfuerzos tendientes a independizar la deducción de las transformaciones de Lorentz de hipótesis vinculadas con fenómenos electromagnéticos. Las demostraciones se fueron depurando progresivamente, desde el planteo inicial debido a Poincaré (1904 y 1905), pasando por los trabajos de Ignatowski (1910), Frank y Rothe (1911), Le Roy (1935), Escanglon (1936) y Lalan (1936) hasta adquirir su forma actual.[72]

Las bases para esta deducción se encuentran en cuatro postulados debidos a Poincaré:
(a) El espacio es homogéneo e isótropo.
(b) El tiempo es homogéneo.
(c) Existen marcos de referencia respecto de los cuales las leyes físicas adoptan la misma forma (marcos inerciales).
(d) Las conexiones causales se conservan al pasar de un marco de referencia inercial a otro.

[72] Ver J. Hladik y M. Chrysos, "*Introduction à la relativité restreinte*", Ellipses, París, 2001.

Actualmente es posible obtener las transformaciones de Lorentz sobre la base de estos cuatro postulados. En lugar de la velocidad de la luz aparece una velocidad límite que no puede ser sobrepasada por los corpúsculos durante su movimiento, y solamente puede ser alcanzada por corpúsculos de masa nula.
Aplicando el principio (c) de relatividad de Poincaré, esa velocidad límite resulta la misma respecto de todos los marcos de referencia inerciales. Es una constante universal que se suele denominar **constante de estructura del espacio-tiempo**.
Si, como pensaba Louis de Broglie, no existieran corpúsculos de masa nula, o si experimentos más precisos mostraran que el postulado de invariancia de la velocidad de la luz en el vacío se puede considerar solamente como una primera aproximación a una realidad más compleja, la deducción de las transformaciones de Lorentz con c sustituida por la mencionada constante de estructura del tiempo-espacio mantendría su validez.

A continuación, a vía de ejemplo, presentamos una breve deducción[73] para el caso en que el eje x' del marco inercial K' se mueve sobre el eje x del marco inercial K, teniendo ambos ejes la misma orientación.
Supongamos que la velocidad con que K' se desplaza respecto de K es V, y que los relojes en cada marco se encuentran sincronizados (por ejemplo trasladando en forma cuasi-estática a diferentes puntos relojes previamente sincronizados en un mismo punto).
En lo que sigue trabajaremos con una sola coordenada espacial. Aunque con alguna complicación adicional, el método se puede extender a tres dimensiones.
Supongamos que las coordenadas de espacio-tiempo de un mismo evento respecto de K y de K' son (x,t) y (x',t') respectivamente.
En general cabe esperar que estas coordenadas se relacionen mediante funciones regulares del tipo: $x' = f(x,t;V)$ $\quad t' = g(x,t;V)$
Considerando eventos próximos y trabajando con diferenciales, se tiene:

$$dx' = \frac{\partial}{\partial x} f(x,t;V)\cdot dx + \frac{\partial}{\partial t} f(x,t;V)\cdot dt \qquad dt' = \frac{\partial}{\partial x} g(x,t;V)\cdot dx + \frac{\partial}{\partial t} g(x,t;V)\cdot dt$$

[73] Esta deducción se basa en el trabajo de Edouard Le Roy, "Sur les formules de Lorentz", Comptes Rendus, vol. 202, 1936, pp. 794-795. Al parecer la demostración se le ocurrió a Le Roy (Profesor en el Collège de France) durante una conversación con Ramón Salas Edwards (Profesor en la Universidad de Chile), mantenida a comienzos de la década de 1930. En los artículos de A. Lee y T. Kalotas "Lorentz transformation from the first principle", American Journal of Physics, vol 43, 1975, pp. 434-437, y de J. Lévy-Leblond "One more derivation of the Lorentz transformation", American Journal of Physics, vol 44, 1976, pp. 271-277, se pueden encontrar otras deducciones de las transformaciones de Lorentz junto con referencias detalladas a los trabajos previos sobre este tema. En el trabajo de J. Ferrari "On the homogeneity of space and time in special relativity", Zeitschrift für Allgemeine Wissenschaftstheorie, vol. 22 (1), 1991, pp. 169-171, puede hallarse una discusión interesante sobre la interdependencia entre la homogeneidad del espacio y la homogeneidad del tiempo en la teoría restringida de la relatividad.

De la homogeneidad del espacio y del tiempo se desprende que las derivadas parciales dependen solamente de V. Cuando $dx' = 0$, entonces $dx = V \cdot dt$ De esta última relación y de la expresión para dx' en función de dx y de dt se desprende que:

$$V \cdot \frac{\partial}{\partial x} f(x,t;V) + \frac{\partial}{\partial t} f(x,t;V) = 0$$

Como consecuencia: $$dx' = \frac{\partial}{\partial x} f(x,t;V) \cdot (dx - V \cdot dt) = \gamma(V) \cdot (dx - V \cdot dt)$$

Como los ejes x y x' poseen la misma orientación, $\gamma(V) > 0$

Si $dx = 0$, entonces: $dx' = -V \cdot dt'$ De esta última relación y de las expresiones para dx' y dt' se obtiene $\frac{\partial}{\partial t} g(x,t;V) = \frac{\partial}{\partial x} f(x,t;V)$ Entonces poniendo $\alpha(V) = \frac{\partial}{\partial x} g(x,t;V)$

resulta: $$dt' = \alpha(V) \cdot dx + \gamma(V) \cdot dt \qquad dx' = \gamma(V) \cdot (dx - V \cdot dt)$$

De estas ecuaciones deducimos que entre la velocidad $u = \frac{dx}{dt}$ de un movimiento de un corpúsculo respecto de K y la velocidad $u' = \frac{dx'}{dt'}$ del movimiento de ese mismo corpúsculo respecto de K' se verifica la relación: $$u' = \gamma(V) \cdot \frac{(u - V)}{(\alpha(V) \cdot u + \gamma(V))}$$

Reordenado: $$\alpha(V) \cdot u' \cdot u + \gamma(V) \cdot (u' - u) = -\gamma(V) \cdot V$$

Si $\alpha(V) = 0$ resulta la ley de adición de velocidades de Galileo: $u' = u - V$

Si $\alpha(V) \neq 0$ existe una velocidad invariante al cambiar de marco de referencia $u' = u = u_0$ que verifica: $\alpha(V) \cdot u_0^2 = -\gamma(V) \cdot V$

Entonces: $$dx' = \gamma(V) \cdot (dx - V \cdot dt) \qquad dt' = \gamma(V) \cdot \left(-\frac{V}{u_0^2} \cdot dx + dt \right)$$

De donde obtenemos: $$dx = \frac{1}{\gamma(V) \cdot \left(1 - \frac{V^2}{u_0^2}\right)} \cdot (dx' + V \cdot dt')$$

Del principio de relatividad de Poincaré se desprenden las relaciones de reciprocidad:

$$dx = \gamma(-V) \cdot (dx' + V \cdot dt') \qquad dt = \gamma(-V) \cdot \left(+\frac{V}{u_0^2} \cdot dx' + dt' \right)$$

De la isotropía del espacio: $\gamma(-V) = \gamma(V)$ Comparando las expresiones para dx y teniendo en cuenta que $\gamma(V)$ debe ser positivo, resulta finalmente: $$\gamma(V) = \frac{1}{\sqrt{1 - \frac{V^2}{u_0^2}}}$$

Del principio de relatividad se deduce que la velocidad u_0 es la misma para todo marco inercial: es una constante de estructura del espacio-tiempo. Además, la fórmula que expresa $\gamma(V)$ en función de u_0 sugiere que esa **velocidad invariante** es al mismo tiempo una **velocidad límite** para todo movimiento de un cuerpo respecto de un marco inercial.

Suponiendo que los orígenes de coordenadas coinciden en $t = t' = 0$ e integrando resultan las fórmulas de transformación de Lorentz:

$$x' = \frac{1}{\sqrt{1 - \frac{V^2}{u_0^2}}} \cdot \left(x - V \cdot t\right) \qquad t' = \frac{1}{\sqrt{1 - \frac{V^2}{u_0^2}}} \cdot \left(-\frac{V}{u_0^2} \cdot x + t\right)$$

7. La interrelación entre el tiempo y el espacio: el continuo de tetra-dimensional de Minkowski y la causalidad en la teoría restringida de la relatividad

"Por lo tanto el espacio por sí mismo, y el tiempo por sí mismo, están condenados a desvanecerse en meras sombras, y solamente una especie de unión de los dos preservará una realidad independiente."

Hermann Minkowski durante una conferencia en 1908, en la Universidad de Göttingen, en la que introdujo el concepto de tiempo-espacio y con ello facilitó recorrer el difícil y tortuoso camino que Einstein, Grossmann y Hilbert emprendieron en busca de una generalización de la teoría restringida de la relatividad capaz de abarcar los fenómenos gravitatorios.

Cada evento E viene caracterizado, respecto de un marco inercial K, por una cuaterna de números reales (t, x, y, z) que corresponden a la coordenada temporal y a las tres coordenadas espaciales. Esto se puede representar así: $(E)_K = (t, x, y, z)$

El mismo evento se puede describir respecto de cualquier marco inercial de referencia. Respecto de otro marco K', un mismo evento E posee otra representación:

$$(E)_{K'} = (t', x', y', z')$$

En la teoría restringida de la relatividad, las coordenadas de espacio y tiempo que representan a un mismo evento respecto de dos marcos inerciales se interrelacionan a través de las fórmulas de Lorentz.

En 6.1 dichas transformaciones y algunas de sus consecuencias fueron analizadas para el caso particular en el cual el eje x' del marco K' se mueve paralelamente al eje x del marco K, coincidiendo los orígenes de los sistemas de coordenadas espaciales en t=t'=0.

Las fórmulas de transformación del instante y de las coordenadas de un mismo evento al cambiar el marco de referencia pueden generalizarse para el caso en que el movimiento relativo de K' respecto de K no se produce necesariamente con los ejes x y x' paralelos entre sí.

Supongamos que $\vec{V}$ es la velocidad de K' respecto de K. Cualquier vector de posición $\vec{r}$ se puede descomponer en la suma de un vector $\frac{(\vec{r} \bullet \vec{V})}{V^2} \cdot \vec{V}$ paralelo a $\vec{V}$ y un vector $\vec{r} - \frac{(\vec{r} \bullet \vec{V})}{V^2} \cdot \vec{V}$ perpendicular a $\vec{V}$.

Aquí $\bullet$ representa el **producto escalar** de vectores y V es la magnitud de $\vec{V}$.

Al pasar al otro marco de referencia la componente $\vec{r} - \frac{(\vec{r} \bullet \vec{V})}{V^2} \cdot \vec{V}$ permanece invariante, mientras que el instante y la componente $\frac{(\vec{r} \bullet \vec{V})}{V^2} \cdot \vec{V}$ se transforman de acuerdo con las fórmulas de Lorentz.

Entonces si $(E)_K = (t, \vec{r})$ es la representación del evento E en el marco K, le representación de ese mismo evento en el marco K', es decir $(E)_{K'} = (t', \vec{r}')$ se conecta con $(E)_K = (t, \vec{r})$ por las expresiones:

$$t' = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}} \cdot \left(t - \frac{\left(\vec{r} \bullet \vec{V}\right)}{c^2} \right) \qquad \vec{r}\,' = \left(\vec{r} - \frac{\left(\vec{r} \bullet \vec{V}\right)}{V^2} \cdot \vec{V} \right) + \frac{1}{\sqrt{1-\frac{V^2}{c^2}}} \cdot \left(\frac{\left(\vec{r} \bullet \vec{V}\right)}{V^2} \cdot \vec{V} - t \cdot \vec{V} \right)$$

Si se define $\gamma = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}}$, la fórmula para transformar el vector de posición y el instante de tiempo se pueden re-escribir así:

$$\vec{r}\,' = \vec{r} + (\gamma - 1) \cdot \frac{\left(\vec{r} \bullet \vec{V}\right)}{V^2} \cdot \vec{V} - \gamma \cdot t \cdot \vec{V} \qquad t' = \gamma \cdot \left(t - \frac{\left(\vec{r} \bullet \vec{V}\right)}{c^2} \right)$$

Parece natural pensar el mundo físico, considerado como un universo de eventos E, recurriendo a un espacio abstracto de cuatro dimensiones.
En 1907 Hermann Minkowski desarrolló esta idea introduciendo un **continuo tiempo-espacio** con una definición de distancia entre dos eventos cualesquiera que se calcula a partir de las coordenadas de espacio y tiempo de los eventos y permanece invariante al cambiar de marco de referencia inercial. La denominó **intervalo** entre eventos.
Si Δt es la separación en el tiempo, y Δx Δy Δz son las separaciones espaciales a lo largo de los ejes de un sistema de coordenadas cartesianas ortogonales, para un par de eventos que se describen respecto de un sistema inercial K, el intervalo Δs entre esos eventos se puede definir así: [74] $(\Delta s)^2 = c^2(\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2$
Respecto de un marco de referencia K', el intervalo entre esos mismos eventos se escribe: $(\Delta s')^2 = c^2(\Delta t')^2 - (\Delta x')^2 - (\Delta y')^2 - (\Delta z')^2$
En forma más compacta, utilizando los correspondientes vectores de posición en el espacio físico se tiene: $(\Delta s)^2 = c^2(\Delta t)^2 - \Delta\vec{r} \bullet \Delta\vec{r}$ $\quad (\Delta s')^2 = c^2(\Delta t')^2 - \Delta\vec{r}\,' \bullet \Delta\vec{r}\,'$
A partir de las fórmulas de transformación de Lorentz se obtiene que $(\Delta s')^2 = (\Delta s)^2$.
El intervalo, esta medida de la distancia entre eventos, en tanto no depende del marco de referencia inercial respecto del cual se la calcula, **es un verdadero absoluto**.
Como el intervalo entre dos eventos conectados por la trayectoria de una señal luminosa vale siempre cero, **el intervalo** no corresponde a una métrica propiamente dicha[75], **es una pseudo-métrica**.

Este punto de vista prioriza la representación del mundo físico mediante un espacio abstracto cuyos puntos corresponden a los eventos.
El movimiento de una partícula material se representa mediante una curva continua de eventos en el espacio de Minkowski, llamada línea de universo.
Desde el punto de vista físico la descripción debe hacerse utilizando algún sistema de coordenadas témporo-espaciales.
El mismo evento se puede describir respecto de cualquiera de estos sistemas de

[74] Como siempre c es la velocidad de la luz en el vacío, o con mayor generalidad, la constante de estructura del espacio-tiempo introducida en 6.3.

[75] Generalmente se pide que una verdadera distancia entre puntos de un espacio abstracto se anule siempre y cuando los puntos coincidan, pero en el caso del intervalo su anulación no asegura la identidad de los sucesos. Por eso se considera una pseudo-métrica.

coordenadas, pero hay que utilizar alguno. Cada sistema de coordenadas representa un marco de referencia inercial con sus relojes sincronizados y sus reglas para medir distancias.

7.1 El grupo de transformaciones de Lorentz y la crono-geometría de Minkowski

Para cada punto E del espacio de Minkowski, o sea, para la representación matemática de cada evento en el mundo físico, las fórmulas de transformación de Lorentz conectan la cuaterna de coordenadas (t, x, y, z) del evento respecto de un sistema de coordenadas K con la cuaterna de coordenadas (t', x', y', z') de ese mismo evento respecto de otro sistema de coordenadas K'.

Consideremos ahora tres sistemas de coordenadas K_1, K_2 y K_3 en el espacio de Minkowski. Representamos por $L_{j,i}$ la transformación de Lorentz que transforma las coordenadas de tiempo-espacio $(E)_{K_i} = (t_i, x_i, y_i, z_i)$ de cada evento E respecto del sistema K_i en sus coordenadas de tiempo-espacio $(E)_{K_j} = (t_j, x_j, y_j, z_j)$ respecto del sistema K_j.

Entonces se puede escribir la transformación entre sistemas de coordenadas, para cada evento E, así: $(E)_{K_j} = L_{j,i}(E)_{K_i}$

En particular: $(E)_{K_2} = L_{2,1}(E)_{K_1}$ $\quad (E)_{K_3} = L_{3,2}(E)_{K_2}$ $\quad (E)_{K_3} = L_{3,1}(E)_{K_1}$

La aplicación sucesiva de $L_{2,1}$ y $L_{3,2}$ se puede escribir $L_{3,2}(L_{2,1}(E)_{K_1})$ y se comprueba que el resultado, válido para todo evento del espacio de Minkowski, es la cuaterna de tiempo-espacio $(E)_{K_3}$ del evento E respecto del sistema K_3. Entonces resulta natural definir una operación de composición $L_{3,2} \circ L_{2,1}$ entre transformaciones de Lorentz tal que $L_{3,2} \circ L_{2,1} = L_{3,1}$.

Cada transformación de Lorentz $L_{j,i}$ posee una transformación inversa $L_{j,i}^{-1} = L_{i,j}$ tal que si I representa la transformación identidad (que deja las coordenadas de cada evento incambiadas), entonces $L_{i,j} \circ L_{j,i} = L_{j,i}^{-1} \circ L_{j,i} = I$

Además, la composición de transformaciones de Lorentz es asociativa:

$$L_{4,3} \circ (L_{3,2} \circ L_{2,1}) = (L_{4,3} \circ L_{3,2}) \circ L_{2,1}$$

Entonces las transformaciones de Lorentz junto con la operación de composición (aplicación sucesiva) de las mismas, forma lo que en álgebra abstracta se conoce como grupo[76].

Como la transformación resultante de esa operación en general depende del orden en que se aplican las transformaciones que se componen, la operación no es conmutativa:

[76] Un grupo G es un conjunto entre cuyos elementos a, b, c,…se ha definido una operación binaria $*$ asociativa (es decir $a * (b * c) = (a * b) * c$ para toda terna de elementos de G), con elemento neutro (identidad) e (es decir tal que $e * a = a * e = a$ para todo elemento a de G) y cada elemento a posee elemento simétrico (inverso) a' (es decir $a * a' = a' * a = e$).

se dice que el grupo así definido es un grupo no conmutativo.
Solo conmutan las transformaciones de Lorentz cuando las velocidades relativas de los marcos de referencia son paralelas.

La teoría de la relatividad puede contemplarse desde dos puntos de vista. En uno de ellos la atención se centra en la relatividad de conceptos como el tiempo y el espacio, que en la física newtoniana se consideraban absolutos. En el otro la atención se vuelve hacia los nuevos absolutos, invariantes bajo el grupo de transformaciones de Lorentz, de los cuales el más básico pero, como se verá en el próximo capítulo, no el único, es el **intervalo**.

Este segundo punto de vista, la del estudio de propiedades de figuras (conjuntos de eventos) que permanecen invariantes bajo un grupo de transformaciones, corresponde al concepto de geometría generalizado por Félix Klein en su famoso Programa de Erlangen.

El análisis de la geometría del continuo tiempo-espacio (lo que puede denominarse **crono-geometría**) utilizando las coordenadas de los eventos respecto de marcos de referencia es una extensión muy eficiente del método analítico de Descartes. Pero, como señala Synge en su exhaustivo tratado de teoría restringida de la relatividad[77], el enfoque analítico a veces parece ocultar los objetos geométricos. "¿Acaso no podemos tratar directamente con esos objetos, como Euclides lo hizo con los puntos, líneas y círculos? Esto es de hecho lo que hizo Robb…Pero la verdad es que estos "métodos sintéticos" son demasiado difíciles."
Por este motivo lo que podría denominarse la crono-geometría sintética de Minkowski ha sido escasamente investigada.

7.2 La relatividad del espacio físico y sus consecuencias. Líneas y tubos de universo.

Una consecuencia de esta interrelación entre el tiempo y el espacio característica del continuo de Minkowski es que la separación de espacio y tiempo, y el orden temporal en el que se presentan algunos eventos dependen del sistema de referencia respecto del cual se establece el orden temporal y espacial de los fenómenos.
En principio, sucesos que son simultáneos respecto de un sistema, no tienen por qué serlo respecto de otro sistema.
No obstante, el ordenamiento temporal de sucesos relacionados mediante una influencia física directa (por ejemplo, dos sucesos vinculados mediante la propagación de una señal) es absoluto, el mismo para todos los sistemas de referencia, de modo que un evento que se puede considerar como causa siempre precede a un evento que se puede considerar como su efecto.

El cuadrado del intervalo entre dos eventos que pueden tener una relación causal es siempre positivo o nulo, porque siendo la velocidad de la luz la máxima velocidad con que pueden viajar las señales, la distancia espacial entre dos eventos causalmente relacionados debe ser inferior o a lo sumo igual a la distancia que viaja la luz durante el intervalo de tiempo que separa los eventos, todo ello respecto de un mismo sistema inercial de referencia.

[77] J. L. Synge "*Relativity: the special theory*", North-Holland, Amsterdam, 2ª edición, 1965, p.58.

Pero los sucesos que no pueden estar relacionados mediante una influencia física directa de uno sobre el otro, presentan un orden temporal dependiente del sistema de referencia respecto del cual son observados. En este caso el cuadrado del intervalo entre los sucesos es siempre negativo (el intervalo es un número imaginario) porque la distancia espacial que separa los eventos es mayor que la distancia que viaja la luz durante el intervalo de tiempo que los separa.

Puesto que el orden espacial en su forma pura se refiere a sucesos simultáneos en lugares diferentes y puesto que la simultaneidad de eventos que no ocurren en exactamente el mismo lugar es relativa, resulta que la separación de las relaciones puramente espaciales a partir del conjunto de relaciones de tiempo-espacio es también relativa, dependiente del sistema inercial utilizado para realizar esa separación.
El **espacio en cuanto tal** es una **sección** del continuo tetra-dimensional tiempo-espacio, sección en la que se encuentran todos los eventos **que son simultáneos respecto de un sistema de referencia dado.**
Pero si cambiamos de sistema de referencia **cambian los eventos integrantes de esa sección, porque son otros los eventos simultáneos.**

Para visualizar esto con facilidad, y puesto que no podemos formarnos imágenes visuales de figuras en cuatro dimensiones, no restringiremos a procesos que se producen en una dimensión espacial. Así podemos representar los eventos en un plano tiempo-espacio, como muestra la Figura 1. Cada punto del plano representa un evento posible.

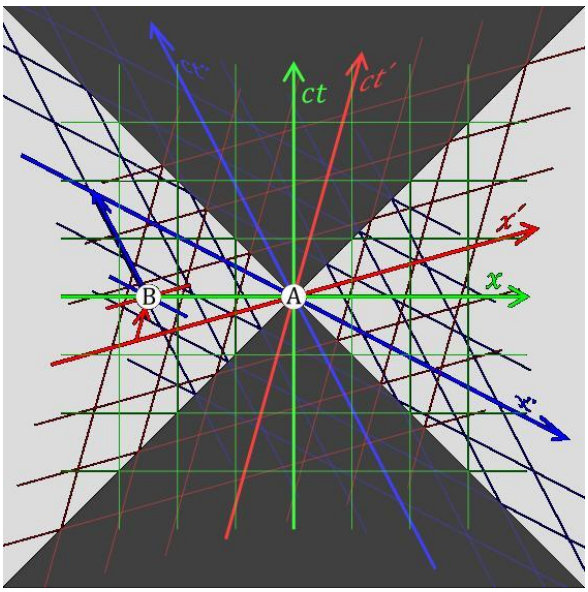


Figura 7.1: Relatividad de la simultaneidad de dos eventos A y B respecto del marco de referencia

Comencemos con un marco K. Representemos con $c \cdot t$ la coordenada temporal[78] y con x la coordenada espacial. **En forma arbitraria** introduzcamos un par de ejes cartesianos ortogonales con el espacio en el eje de abscisas y el tiempo en ordenadas, como se ve en la figura.
Ahora introduzcamos un nuevo marco K' cuyo eje x' se desplaza con velocidad

[78] Al multiplicar por la velocidad de la luz se obtiene una medida del tiempo con dimensiones de espacio.

positiva V a lo largo del eje x de K.

Definiendo $\beta = \frac{V}{c}$ y $\gamma = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}}$ las fórmulas de transformación de Lorentz se pueden re-escribir así: $x' = \gamma \cdot (x - \beta . c \cdot t)$ $c \cdot t' = \gamma \cdot (c \cdot t - \beta \cdot x)$

El eje x' corresponde a $c \cdot t' = 0$, es decir a la recta $c \cdot t = \beta \cdot x$ respecto de los ejes cartesianos ortogonales. El eje $c \cdot t'$ corresponde a $x' = 0$, o sea a la recta $c \cdot t = \frac{1}{\beta} \cdot x$ respecto de los ejes del marco K. Como consecuencia los ejes de coordenadas temporal y espacial del marco K' forman un ángulo agudo entre sí, y quedan situados en el primer cuadrante de K como muestra la figura.

Finalmente, para un tercer marco K'' que se mueve en sentido opuesto a K', es decir con $\beta = \frac{V}{c} < 0$, los correspondientes forman un ángulo obtuso y se disponen como muestra la figura.

Como los ejes de coordenadas de los marcos K' y K'' son oblicuos, las coordenadas de un evento respecto de esos marcos se obtienen por desplazamiento paralelo a uno de los ejes, hasta cortar el otro, y no por proyección ortogonal como en el caso del marco K.

En la Figura 1 se marcaron dos eventos, el A situado en el origen común de coordenadas, y el B que es simultáneo con A respecto de K, posterior a A respecto de K' y previo a A respecto de K''. (Estas relaciones se indican mediante flechas que llegan a B desde, o parten de B hacia, el eje formado por los eventos simultáneos a A en los marcos K' y K'' respectivamente).

Cuando V es positiva, como $0 \leq \beta = \frac{V}{c} \leq 1$, a medida que $\beta = \frac{V}{c}$ crece el ángulo entre el eje $c \cdot t'$ y el eje x' tiende a cero, hasta que en el límite, cuando V se aproxima por debajo a c, ambos ejes se superponen con la bisectriz del primer cuadrante del plano de eventos referido al marco K. Esta observación resulta de importancia en relación con el cono de luz y las conexiones causales en teoría restringida de la relatividad como se verá posteriormente.

En suma, si una sección espacial del continuo tiempo-espacio es el conjunto de eventos simultáneos con un evento dado respecto de un marco de referencia determinado, la Figura 1 muestra que esta sección espacial es diferente para marcos diferentes (consiste en los eventos ubicados sobre los ejes de coordenadas espaciales y sus rectas paralelas).

Por ser el espacio físico una sección del continuo tiempo-espacio asociada al mismo instante respecto de un marco de referencia, no hay tiempo cósmico ni espacio cósmico que puedan considerarse con independencia de un marco: cambian al cambiar el marco. Por tanto, de acuerdo con la teoría restringida de la relatividad, **no es posible pensar en el estado del Universo en un instante de tiempo**, como pensaban los astrónomos y cosmólogos presuponiendo un tiempo absoluto separado de un espacio absoluto, antes de la aparición de la física relativista[79].

[79] No obstante, como se verá posteriormente, al estudiar el problema cosmológico, el postulado de Weyl introduce una clase de observadores privilegiados cada uno de cuyos tiempos propios se pueden correlacionar con los tiempos propios de los de los demás observadores de tal forma que se puede introducir un **tiempo cósmico universal.**

Si un cuerpo se mueve respecto de otro cuerpo, no es posible hablar de la distancia entre ambos cuerpos en un instante de tiempo dado sin especificar un marco de referencia. Como consecuencia el enunciado de la ley de gravitación de Newton es ambiguo y no resulta aceptable desde la perspectiva de la física relativista, puesto que involucra una distancia entre cuerpos que se supone independiente del marco de referencia.
En la geometría física también aparecen problemas. Por ejemplo, una línea recta en el espacio físico es una cierta traza cuyas partes componentes se manifiestan en forma simultánea para un observador (o una placa fotográfica). Pero como la simultaneidad es relativa, respecto de otro observador esas mismas partes podrán no aparecer ya como simultáneas, así que para este nuevo observador no será una línea recta.

A cada sistema material con un volumen finito le corresponde un **tubo de universo** incluido en el tiempo-espacio de Minkowski: una figura geométrica en dicho espacio tetra-dimensional. Si se trata de una partícula, le corresponde una curva en el tiempo-espacio denominada **línea de universo**.
Respecto de un marco de referencia K, el tubo de universo de un sistema material puede seccionarse mediante hiper-planos t=constante y surgen así regiones espaciales tridimensionales cuyos puntos corresponden a eventos simultáneos respecto de K: el volumen que ocupa cada una de ellas es el volumen del sistema en ese instante de tiempo referido a K.
De esta forma un observador en reposo respecto de K define el volumen que ocupa el sistema en un instante de tiempo y describe sus variaciones temporales.
Pero si utilizamos otros marcos inerciales K', K'',…, seccionamos el tiempo-espacio por los hiper-planos t'=constante, t''=constante,…, con lo cual las regiones espaciales tridimensionales ocupadas por eventos pertenecientes al cuerpo estudiado están formadas por eventos diferentes y los volúmenes del sistema son ahora los volúmenes de estas regiones.
Entonces un observador en reposo respecto de uno de estos marcos, digamos respecto de K', determina el volumen de ese sistema, en lo que para él es un mismo instante de tiempo, involucrando un conjunto de eventos diferente al involucrado cuando el observador en K realiza su medición.

Como consecuencia esos conjuntos de eventos, los referidos a K y los referidos a K', **no se relacionan entre sí** mediante transformaciones de Lorentz.
Fermi se había dado cuenta y había analizado el problema ya en 1923[80], durante un cálculo de la masa electromagnética mediante un modelo relativista de un electrón considerado como una carga eléctrica distribuida en una región cuya dimensión característica es el radio clásico atribuido a esa partícula y visto en 4.3.

7.3 El tiempo en la teoría restringida de la relatividad

Se ha dicho, y se dice a veces, que la teoría de la relatividad espacializa el tiempo. Este punto de vista ha sido sostenido casi siempre por personas inclinadas a pensar que el tiempo es en última instancia una ilusión.
La posibilidad de concebir el continuo de Minkowski como algo real y dado en su totalidad, que puede ser visto de afuera abarcando desde el pasado más remoto hasta el futuro más lejano, para todo marco de referencia posible, resulta atractiva para personas

[80] A. Gamba "Physical quantities in different reference systems according to relativity", American Journal of Physics, Vol. 35, 1967, pp.83-89.

que piensan en un universo material causalmente cerrado y estrictamente determinista.

En la medida en que la teoría de la relatividad nos conduce a considerar el conjunto de todos los sucesos, haciendo abstracción de sus propiedades concretas, atendiendo solo a la estructura universal de relaciones de acción causal entre ellos, por la cual todo suceso del conjunto actúa, de una u otra manera, sobre otros sucesos y recibe a su vez la acción de los otros, **ese mundo de eventos que de ese proceso de abstracción resulta presenta tiempo y espacio fundidos en un continuo**.
Al intentar caracterizar un evento, es necesario considerar un sistema de coordenadas en ese continuo. Cada sistema de coordenadas admisible para la teoría restringida de la relatividad corresponde a un sistema inercial con sus relojes y sus reglas. La geometría correspondiente a ese continuo se puede caracterizar por las configuraciones que permanecen invariantes bajo el grupo de todas las transformaciones entre esos sistemas de coordenadas (que no son otras que las transformaciones de Lorentz).

La pseudo-métrica que se introduce a través del intervalo que separa dos eventos, conlleva la multiplicación del intervalo temporal por la velocidad de la luz, lo cual equivale a sustituir la distancia en el tiempo por una distancia espacial equivalente (la distancia que recorrería una señal luminosa durante ese tiempo), de modo que las cuatro coordenadas de cada evento son espaciales.
Claro está que se podría haber dividido las coordenadas espaciales por la velocidad de la luz, y entonces, como las nuevas coordenadas miden el tiempo que una señal luminosa tardaría en recorrer los correspondientes intervalos espaciales, ahora las cuatro coordenadas de un evento serían temporales.

Ambas descripciones son equivalentes, así que optar por una u otra es un asunto de gusto o más bien de predisposición.
En la descripción en términos de coordenadas temporales el invariante que mide la distancia entre los eventos es la diferencia de **tiempo propio** entre ellos.
La extensión $\tau_2 - \tau_1$ del intervalo de tiempo propio que mide un observador fijo a un marco inercial que se mueve con velocidad V respecto de otro marco inercial se vincula con la extensión $t_2 - t_1$ del intervalo de tiempo correspondiente medido por el otro observador por la fórmula: $\tau_2 - \tau_1 = \sqrt{1 - \frac{V^2}{c^2}} \cdot (t_2 - t_1)$
Consideremos ahora una partícula en movimiento, que desde el punto de vista geométrico recorre su línea de universo en el tiempo-espacio de Minkowski.
Para caracterizar el movimiento es necesario referirlo a un marco inercial K. Supongamos que la partícula se mueve con velocidad variable respecto de un observador fijo en K.
Si la trayectoria de una partícula respecto de un marco K viene dada por $x = x(t)$ $y = y(t)$ $z = z(t)$ entonces el intervalo $(\Delta s)^2 = c^2(\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2$ entre dos eventos sobre que corresponden al paso de la partícula por dos puntos sucesivos de su trayectoria puede re-escribirse como un elemento diferencial sobre esa trayectoria:

$$ds = \sqrt{c^2 - (\frac{dx(t)}{dt})^2 - (\frac{dy(t)}{dt})^2 - (\frac{dz(t)}{dt})^2} \cdot dt = c \cdot \sqrt{1 - \frac{v(t)^2}{c^2}} \cdot dt = c \cdot d\tau(t)$$

En estas igualdades $v(t)=\sqrt{(\frac{dx(t)}{dt})^2+(\frac{dy(t)}{dt})^2+(\frac{dz(t)}{dt})^2}$ es la magnitud de la velocidad (en general variable) de la partícula respecto de K.
Consideremos ahora su integral sobre la trayectoria de la partícula entre dos instantes t_1 y t_2 registrados en un reloj sincronizado en K:

$$\tau_2-\tau_1=\int_{t_1}^{t_2}\sqrt{1-\frac{v^2(t)}{c^2}}\cdot dt$$

De esta última ecuación, teniendo en cuenta que $\sqrt{1-\frac{v^2}{c^2}}$ es menor que 1 cuando $v(t)\neq 0$, se deduce que el intervalo de tiempo $\tau_2-\tau_1$ es siempre más corto que el correspondiente intervalo de tiempo registrado en un marco de referencia inercial respecto del cual se describe el movimiento de la partícula.
¿Qué significado posee $\tau_2-\tau_1$? Generalmente se supone que $\tau_2-\tau_1$ es la extensión de un intervalo de tiempo propio, es decir, el tiempo tal como sería medido por un reloj fijo a la partícula. Se admite entonces que la fórmula $d\tau(t)=\sqrt{1-\frac{v^2(t)}{c^2}}\cdot dt$ se puede utilizar para relacionar el tiempo medido por un observador acelerado con el tiempo medido por un observador fijo a un marco de referencia inercial.
No obstante, cuando la partícula se encuentra acelerada respecto del marco inercial K, un marco de referencia N solidario con ella no es ya un marco inercial. Entonces **no hay forma de conectarlo con K mediante una única transformación de Lorentz**.
En la teoría restringida de la relatividad, un observador fijo a un marco inercial puede describir en términos de tiempo y espacio un movimiento cualquiera de un cuerpo o sistema de cuerpos, incluyendo por tanto el movimiento de un observador no inercial. Pero este último no puede hacer una descripción equivalente a la del observador inercial, a menos que abandonemos el marco conceptual propio de la relatividad restringida y pasemos al de la teoría generalizada.
Como consecuencia la interpretación de $d\tau(t)=\sqrt{1-\frac{v^2(t)}{c^2}}\cdot dt$ como elemento de tiempo propio debería ser considerada como un nuevo postulado de la teoría restringida. Sus límites de aplicabilidad deberían discutirse y establecerse en la teoría generalizada, pero no se puede descartar a priori una limitación en las aceleraciones permisibles[81].

7.4 Causalidad y determinismo en el marco de la teoría de relatividad

En el proceso de integración de tiempo y espacio en un mismo continuo de cuatro dimensiones, el espacio parece llevar la peor parte, al menos si se toma como punto de comparación el concepto newtoniano absoluto.
Como lo señaló Karl Popper[82], en la teoría restringida de la relatividad, pese a la relatividad de la simultaneidad, (que tanto afecta al concepto de espacio en cuanto tal), lo esencial del concepto de tiempo, el antes y después, su relación con la causalidad,

[81] H. Schwarz "*Introduction to special relativity*" Mc Graw-Hill, New York, 1968, p. 63.

[82] En "*El Universo Abierto: Un argumento a favor del indeterminismo*", Tecnos, Madrid, 1996, pp.80-85.

con la asimetría entre el pasado y el futuro, reaparece abriendo nuevas posibilidades epistemológicas.

La diferencia de signo que presenta la componente temporal respecto de las componentes espaciales del intervalo, corresponde a una asimetría en los roles del tiempo y el espacio.
Mientras que se puede ir de derecha a izquierda y de izquierda a derecha a lo largo de un eje de coordenadas espaciales, en el continuo tiempo-espacio de la teoría restringida de la relatividad se puede influir físicamente sobre eventos futuros pero no sobre sucesos pasados.

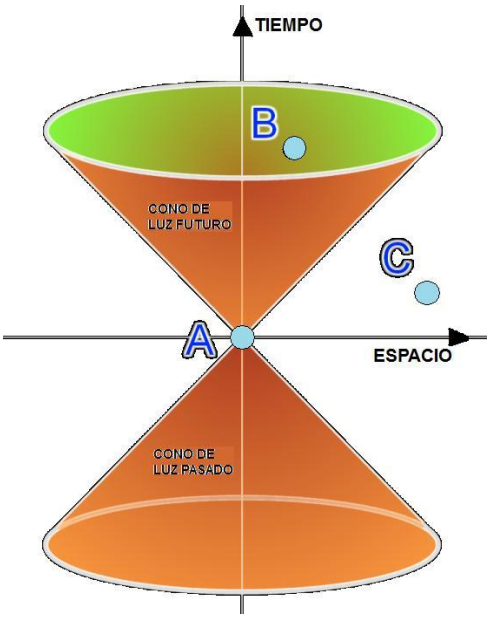


Figura 7.2 El cono de luz según Arthur Eddington

Arthur Eddington[83] propuso representar geométricamente las relaciones entre tiempo y espacio mediante un diagrama que se hizo famoso: la representación de los conos de luz. Limitó las dimensiones del espacio a dos, de modo que añadiendo una dimensión correspondiente al tiempo, se puede hacer el conocido dibujo en perspectiva que aparece en la mayor parte de las obras didácticas y se muestra en la Figura 2.
En lo que sigue se efectúa una descripción verbal abstracta de esos conos. Tomando un evento A (**aquí-ahora**) como origen y un sistema inercial como base para determinar las coordenadas de tiempo y espacio, la totalidad de los puntos (eventos) del continuo de Minkowski se reparten en tres regiones, separadas entre sí por la superficie de los denominados **conos de luz pasado y futuro**:

$$c^2 \cdot t^2 - x^2 - y^2 - z^2 = 0$$

El eje de coordenadas de tiempo es el eje común a ambos conos, que tienen en el evento tomado como origen su vértice común.
La magnitud Δs^2 toma valores positivos en el interior de los conos, negativos en el interior de la región del tiempo espacio que rodea a los conos y cero en la superficie de los mismos.
Los eventos tales como B, que se encuentran en el interior del cono superior en la

[83] En "*The Nature of the Physical World*", Cambridge, UK, 1928.

Figura 2, se encuentran siempre en el futuro respecto de A (de ahí la designación de cono de futuro absoluto), mientras que los eventos en el interior del cono inferior se encuentran siempre en el pasado respecto de A (de ahí la designación de cono pasado absoluto).
La región que rodea a los conos se puede denominar **presente posible** (visto desde la perspectiva del tiempo, porque cualquier evento allí situado, tal como el C, puede hacerse simultáneo con el origen si se elige adecuadamente el sistema de referencia) o **más allá absoluto** (visto desde la perspectiva del espacio, porque no es posible encontrar un sistema de referencia en el que uno de estos eventos se produzca en la misma posición espacial que el origen).
La dirección de la línea de universo de una partícula con masa de reposo no nula, se encuentra en cada uno de sus puntos en el interior del cono de luz correspondiente al punto. La dirección de la línea de universo de los corpúsculos con masa de reposo nula, como los fotones, se encuentra en cada uno de sus puntos en la superficie del correspondiente cono de luz.

Si la velocidad de la luz se hace tender a infinito, el cono futuro se amplía hasta abarcar el hemi-espacio $t \geq 0$, mientras que el cono pasado se amplía hasta abarcar el hemiespacio $t \leq 0$.
El hiperplano $t = 0$ que los separa, representa ahora el presente posible, el mismo para todos los sistemas de referencia inerciales: es el **presente absoluto** de la física newtoniana, formado por eventos que son simultáneos para cualquier observador.

A partir de los conos de luz y su significado causal, es posible concluir que la versión de Laplace del determinismo científico no resulta sustentable desde la perspectiva de la teoría restringida de la relatividad.
Supongamos que la predicción de los estados futuros de un sistema aislado se efectúa a partir de toda la información que en principio podría juntar un observador situado aquí y ahora.
Ese observador conoce las leyes deterministas de la dinámica y dispone de los medios de cálculo necesarios.
Llamemos O al evento asociado con el observador. Sea F un evento futuro que nuestro observador pretende predecir. Por ser un evento futuro se encuentra en el cono futuro de O. Si un evento, tal como F, es posterior a otro, tal como O, es decir si F se encuentra en el cono futuro de O, que entonces le es previo, el cono pasado de F abarca completamente el cono pasado de O. Pero **el cono pasado del evento posterior F comprende eventos adicionales respecto al cono pasado del evento previo O**, de modo que hay sucesos que pertenecen al cono pasado de F pero no pertenecen al cono pasado de O. Como estos eventos pueden influir sobre el suceso que va a ocurrir en F, pero O no puede recibir información sobre ellos porque no se encuentran en el cono pasado de O, el observador no dispone de suficiente información como para poder predecir lo que va a ocurrir en F.
Parecería entonces que la teoría restringida de la relatividad conduce a considerar nuestro pasado como aquella región del continuo tiempo-espacio que nosotros, aquí y ahora, podemos en principio conocer, mientras que nuestro futuro es aquella región que, sometida a la influencia del pasado, está siempre abierta y en principio no es plenamente cognoscible.

Lo que la teoría restringida permite son retrodicciones completas, no predicciones

completas. Eso a pesar de que sus leyes dinámicas son deterministas.

En todo esto presuponemos que el observador que predice o retrodice se encuentra en el interior del sistema. Las leyes disponibles son deterministas en el sentido que si se las alimenta de información completa y exacta sobre las condiciones iniciales de un sistema aislado, permiten predecir o retrodecir los estados del sistema con exactitud.

El retardo en la recolección de la información proveniente de los puntos distantes genera en última instancia esta asimetría entre pasado y futuro. Si se pudiera recolectar información sin retardo alguno desde cualquier punto del espacio, como presupone la formulación de Laplace sobre la posibilidad de predecir con absoluta precisión los estados futuros de un sistema aislado (el universo en su caso), se podría definir un tiempo absoluto. Entonces, como dijimos antes, los conos de luz degenerarían en semi-espacios separados por un hiper-plano de presente absoluto. Conociendo los sucesos simultáneos en ese presente absoluto (el estado presente del sistema) y las leyes dinámicas deterministas, se podrían predecir los estados futuros o reconstruir los estados pasados con absoluta precisión. Una vez aceptada la teoría restringida de la relatividad, esta posibilidad quedó vedada.

8. Física relativista

"De todo lo expuesto se deduce que el principio de relatividad es equivalente a la afirmación de que las leyes de la naturaleza tienen que verificarse cuando las variables de espacio y tiempo sufren una transformación de Lorentz. Por ello se dice que ellas son relativistamente invariantes. La Física debe presentar claramente sus leyes en formas que posean tales invariancias; cualquier principio o enunciado que falle a este respecto no puede ser una ley general."
W. Mc Crea, "Física relativista", UTHEA, México, 1965.

Durante los siglos XVII y XVIII se fue consolidando, en torno a la mecánica newtoniana, la idea de una única estructura matemática subyacente a los fenómenos del mundo físico. Se pensaba en esa época, e inclusive todavía lo pensaban muchas personas durante los siglos XIX y XX, que esa estructura es accesible al entendimiento humano y se expresa en las fórmulas matemáticas que los científicos irían descubriendo en forma progresiva y acumulativa.
La aparición y posterior desarrollo de la física relativista originó la aparición de ramas de la física paralelas a bien desarrolladas ramas de la física clásica, como la dinámica de partículas, la mecánica de medios continuos, la óptica física e inclusive la termodinámica.
No obstante las nuevas disciplinas contenían ciertos resultados incompatibles con algunos de los contenidos de las correspondientes ramas clásicas de la física, excepto en un caso: cuando las velocidades en juego eran pequeñas respecto de la velocidad de la luz. Si esto ocurría, las nuevas fórmulas se reducían siempre a las antiguas.
Pero, como se desprende de los conceptos relativistas de tiempo y espacio estudiados en el capítulo 7, el tiempo que aparece como parámetro en la dinámica de Newton no es compatible con el tiempo que aparece en las ecuaciones de la física relativista, pese a que estas últimas se reduzcan a las primeras cuando, al tender la velocidad de la luz a infinito, los conos de luz de los eventos se abran hasta abarcar dos semi-espacios completos.

Dos décadas después de la aparición y consolidación de la física relativista, se consolidó la física cuántica en los trabajos de Schrödinger, Heisenberg y Dirac, con consecuencias aún más revolucionarias y destructoras para la visión newtoniana del mundo.
Por otra parte, las ideas acerca de la naturaleza de la matemática y su relación con el mundo fenoménico se modificaron profundamente durante el siglo XIX y el primer tercio del siglo XX. El ideal de una estructura cerrada con un fundamento definitivo, perseguido por los formalistas encabezados por Hilbert, debió ceder frente a la idea de estructura abierta que se desarrolló a partir de los resultados de Gödel y Lakatos.

Estos cambios en los conceptos sobre la matemática y la física, condujeron a que la antigua idea de que existe una única estructura matemática profunda, subyacente a la realidad y susceptible de ser descubierta durante un proceso de progreso ininterrumpido, fuera paulatinamente abandonada por la mayoría de los investigadores en favor de la idea de que la estructura matemática de la física teórica consiste en modelos que se construyen para ayudar a comprender el mundo físico: explicar en lo posible las relaciones entre fenómenos ya conocidas y predecir otras nuevas.
En la medida en que solo son modelos matemáticos, resulta admisible que puedan presentar ciertas incompatibilidades.

Como señaló Bridgman[84], luego de la consolidación de las físicas relativista y cuántica, parece poco plausible que nuevas revoluciones científicas modifiquen sustancialmente las ideas sobre el rol del observador en física. Pero debemos reconocer que la lógica, las matemáticas y la física teórica son en buena medida inventos nuestros, y que por ello no es prudente esperar de ellas un éxito rotundo y definitivo.

Antes de comenzar a analizar este tema, al final del presente capítulo, revisaremos algunos aspectos de física relativista. Los aspectos seleccionados, además de su importancia intrínseca en el marco de la teoría restringida, son necesarios para un abordaje histórico de la teoría generalizada en los capítulos 10 y 11, y su aplicación a la denominada cosmología científica y a la astrofísica estelar en los capítulos 12 y 13.
Parte del contenido de las secciones 8.5 (nociones sobre cuadrivectores y campo electromagnético) y 8.7 (nociones sobre termo-mecánica de fluidos relativista) es un poco más abstracto, y para algunos lectores podría resultar algo más demandante desde el punto de vista matemático, que el resto del capítulo. Pero para continuar con la lectura del libro alcanza con captar las ideas principales allí expuestas.

8.1 Dinámica relativista y leyes de conservación

Para satisfacer el **principio de relatividad restringida**, las ecuaciones de la dinámica deben formularse de tal modo que adquieran la misma forma respecto de cualquier marco inercial: esto se resume diciendo que deben ser covariantes bajo el grupo de transformaciones de Lorentz.
En este capítulo seguiremos un orden de exposición parecido al histórico. Por ese motivo, recién en la parte 8.5 consideramos el concepto de cuadrivector asociado al espacio de Minkowski y cómo este concepto permite una formulación matemática precisa del requerimiento de covariancia.

Además, la formulación de la dinámica relativista debe satisfacer el **principio de consistencia** según el cual una nueva teoría (en este caso la mecánica basada en la teoría restringida de la relatividad) que pretende reemplazar a una más antigua (la mecánica de Newton) debe dar cuenta de todas las predicciones exitosas de esta última.

8.1.1 Dinámica de una partícula

Definiendo la cantidad de movimiento o momento lineal $\vec{p}$ de una partícula como el producto de su masa inercial m por su velocidad $\vec{v}$, la segunda ley de Newton se puede escribir de esta forma $\frac{d}{dt}\vec{p} = \vec{F}$.

Como la masa de la partícula se supone constante $\frac{d}{dt}\vec{p} = m \cdot \frac{d}{dt}\vec{v} = m \cdot \vec{a}$, con lo que se recupera la formulación usual de la segunda ley: $m \cdot \vec{a} = \vec{F}$
Puesto que en su forma original la segunda ley es invariante bajo transformaciones de Galileo, para poderla extender al caso relativista es necesario formularla de modo que resulte invariante bajo las transformaciones de Lorentz.

[84] W. Dampier "Historia de la ciencia y sus relaciones con la filosofía y la religión", Tecnos, Madrid, 1997, p. 518.

Puesto que las leyes de la electrodinámica clásica son invariantes bajo estas últimas transformaciones, el problema de obtener una formulación relativista para la dinámica se encaró originalmente desde la electrodinámica.
No obstante es posible obtener la solución desde una perspectiva puramente mecánica, definiendo la cantidad de movimiento y la energía relativista de modo tal que cumplan las correspondientes leyes de conservación y se reduzcan a sus expresiones clásicas cuando las velocidades de las partículas sean despreciables respecto de c.

Si admitimos la cinemática relativista, mantenemos la definición de cantidad de movimiento como producto de masa por velocidad, e imponemos la conservación de la cantidad de movimiento del sistema durante el choque elástico de dos esferas de igual masa (cuando se las mide respecto del mismo marco de referencia), es posible demostrar, siguiendo a Lewis y Tolman[85], que la masa que puede denominarse masa relativista debe necesariamente modificarse en función de la velocidad según la ley general:

$$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$$

Cuando la velocidad de la partícula resulta despreciable respecto de la velocidad de la luz, esta fórmula se reduce a $m = m_0$, siendo m_0 homologable a la masa inercial de la mecánica newtoniana. En física relativista se la conoce como **masa en reposo,** o simplemente **masa** y se asume que es **invariante** bajo las transformaciones de Lorentz[86].

Entonces la cantidad de movimiento viene dada por $\vec{p} = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} \cdot \vec{v}$

Con la nueva definición de cantidad de movimiento la extensión relativista de la segunda ley es inmediata: $$\frac{d}{dt}\left(\frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} \cdot \vec{v}\right) = \frac{d}{dt}\vec{p} = \vec{F}$$

Pero como en dinámica relativista se postula esta formulación matemática de la ley del movimiento, su validez debe confirmarse confrontando sus consecuencias operacionales con los resultados de experimentos adecuados.
La comprobación experimental ha sido realizada empleando fuerzas de origen electromagnético, calculables mediante la fórmula de Lorentz $\vec{F} = q \cdot \left(\vec{E} + \vec{v} \times \vec{B}\right)$ cuyas componentes eléctrica y magnética fueron explicadas en la parte 4.1.

---

[85] Ver G. Lewis y R. Tolman "The principle of relativity and non-newtonian mechanics", Proceedings of the American Academy of Arts and Sciences, vol. 44, Nº 25, 1909, pp. 711-724, o bien R. Tolman "*The theory of relativity of motion*", University of California Press, Berkeley, 1917, pp. 40-42. Una deducción simple de la fórmula para la masa relativista se encuentra en el capítulo 3 del libro de R. Resnick "*Conceptos de relatividad y física cuántica*", Limusa, México, 1976.

[86] Que la masa en reposo sea un invariante relativista no significa que permanezca inmutable: puede desaparecer parcial o totalmente, como se indica más adelante en esta misma sección.

Debido a la aparición del factor $\frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$ de masa relativista (que varía si varía la magnitud de la velocidad) en lugar del factor de masa newtoniana, en general la fuerza relativista no es paralela a la aceleración, a diferencia de lo que acontece con la fuerza y la aceleración en mecánica newtoniana.

Si el movimiento respecto de un marco K se produce a lo largo de una recta, la fuerza es paralela a la aceleración y la fórmula del movimiento se reduce a la expresión escalar:

$$\frac{d}{dt}\left(\frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}\cdot v\right)=\frac{m_0}{\left(1-\frac{v^2}{c^2}\right)^{\frac{3}{2}}}\cdot\frac{dv}{dt}=F$$

Si se define una masa longitudinal $m_l=\frac{m_0}{\left(1-\frac{v^2}{c^2}\right)^{\frac{3}{2}}}$, resulta que la expresión relativista para la segunda ley adopta una forma análoga a la de la mecánica clásica $m_l\cdot a=F$, pero con una masa longitudinal que aumenta al aumentar la velocidad del cuerpo. Cuando el movimiento no es unidimensional, se introduce una masa transversal

$$m_t=\frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}=m\,.$$

Cuando las relaciones entre la velocidad y estas dos masas relativistas, (masa longitudinal y masa transversal) se comparan con los resultados experimentales como los obtenidos por Kaufmann (1902) y Bucherer (1909), y con las mediciones realizadas en aceleradores de electrones, el acuerdo es excelente. La dinámica relativista de las partículas cargadas en campos electromagnéticos resulta una herramienta fundamental para diseñar aceleradores para uso médico, industrial o en investigación[87]. Si no se tiene en cuenta el aumento relativista de la masa en el diseño del acelerador, éste no funciona.

La variación ΔE en la energía del cuerpo, cuando su velocidad aumenta desde cero en el instante inicial hasta u en el instante t, se puede definir como la integral de la potencia $F\cdot v$:

$$\Delta E=\int_0^t F(t)\cdot v(t)\cdot dt=\int_0^t\frac{m_0}{\left(1-\frac{v^2}{c^2}\right)^{\frac{3}{2}}}\cdot v\cdot\frac{dv}{dt}\cdot dt=\frac{m_0\cdot c^2}{\left(\sqrt{1-\frac{v^2}{c^2}}\right)}-m_0\cdot c^2=m\cdot c^2-m_0\cdot c^2$$

Así pues la variación de energía se relaciona con un cambio en la masa relativista:

$$\Delta E=c^2\cdot\Delta m$$

[87] M. S. Livingston "*High-energy accelerators*", Interscience, N. Y., 1954; E. Wilson "*Particle accelerators*", Oxford University Press, N. Y., 2001; H. Wiedemann "*Particle accelerator physics*", Springer, N. Y., 2007.

Cuando la velocidad de la partícula es despreciable respecto de c, la variación en la energía del cuerpo se reduce a la energía cinética de la mecánica clásica:

$$\Delta E = \frac{1}{2} \cdot m_0 \cdot v^2$$

En general ΔE representa la energía cinética relativista de un cuerpo de masa en reposo m_0. Si la energía total de un cuerpo que se mueve libremente respecto de un marco de referencia es $\frac{m_0 \cdot c^2}{\left(\sqrt{1-\frac{v^2}{c^2}}\right)}$, y si ΔE es su energía cinética, entonces $m_0 \cdot c^2$ se puede interpretar como una energía asociada a la masa reposo del cuerpo. Es posible justificar la expresión de la energía para un cuerpo en movimiento libre y su conservación para un sistema de partículas.

Cuando v se aproxima por debajo a la velocidad de la luz en el vacío, la energía y la cantidad de movimiento tienden a infinito, al igual que el trabajo que es preciso realizar para acelerar el cuerpo. Esto significa que **la velocidad de la luz no es alcanzable por un cuerpo cuya masa en reposo sea diferente de cero**.

Eliminando la velocidad de la partícula partir de las relaciones $E = \frac{m_0 \cdot c^2}{\left(\sqrt{1-\frac{v^2}{c^2}}\right)}$ y $p = \frac{m_0 \cdot v}{\left(\sqrt{1-\frac{v^2}{c^2}}\right)}$ se obtiene la fórmula siguiente, que relaciona la energía con la cantidad de movimiento:

$$E = \sqrt{m_0^2 \cdot c^4 + c^2 \cdot p^2}$$

8.1.2 Sistemas de partículas y leyes de conservación

Cuando se considera un sistema aislado de partículas que han interactuado entre sí durante un tiempo y luego se mueven libremente, tanto en mecánica clásica como en mecánica relativista se verifican ciertas leyes de conservación en forma independiente del tipo de interacción. Estas leyes de conservación, de considerable interés, no solo físico sino también filosófico, se pueden obtener a partir de las propiedades de homogeneidad del tiempo y de la homogeneidad e isotropía del espacio[88].

En ausencia de un campo de fuerzas externo al sistema, o en presencia de un campo de fuerzas estacionario y conservativo, las propiedades del sistema como un todo no cambian con el paso del tiempo debido a que todos los instantes son equivalentes con respecto de ese sistema. Una consecuencia de esto es que la energía total del sistema se mantiene constante durante el movimiento.

---

[88] La formulación de la mecánica, tanto clásica como relativista, a partir del principio de Hamilton de mínima acción, permite la deducción más clara y directa de las leyes de conservación asociadas a estas simetrías del tiempo y del espacio. Ver por ejemplo, los libros de L. Landau y E. Lifchitz "*Mecánica*" y "*Teoría del campo*", Reverté, Barcelona, 1982.

Por ejemplo, supongamos que un cuerpo libre de masa en reposo m_0 se encuentra inmóvil respecto de un marco K y que se fragmenta espontáneamente en dos cuerpos libres de masas en reposo m_1 y m_2. Esos cuerpos poseen velocidades $\vec{v}_1$ y $\vec{v}_2$ respectivamente, referidos al marco K.
La conservación relativista de la energía da para este caso:

$$m_0 \cdot c^2 = \frac{m_1 \cdot c^2}{\sqrt{1-\frac{v_1^{\,2}}{c^2}}} + \frac{m_2 \cdot c^2}{\sqrt{1-\frac{v_2^{\,2}}{c^2}}}$$

Como los denominadores son siempre menores que 1, esta igualdad exige que la masa en reposo del cuerpo sea mayor que la suma de las masas en reposo de los fragmentos: $m_0 > m_1 + m_2$ En forma equivalente el defecto de masa $\Delta m = m_0 - m_1 - m_2$ debe ser positivo. La energía asociada con la masa en reposo Δm que desaparece, aparece en forma de energía cinética de los fragmentos. Si estos se frenan al atravesar un material, su energía cinética se transforma en calor, como ocurre con los fragmentos de fisión de un núcleo de uranio o plutonio en el combustible de un reactor nuclear[89].
Si el defecto de masa Δm fuera negativo, la fragmentación espontánea no sería posible porque violaría la conservación de la energía. No obstante, se lo podría fragmentar inyectándolo energía desde el exterior, en cantidad por lo menos igual a $-\Delta m \cdot c^2$.

La relación entre la masa en reposo y la energía de un cuerpo permite cuantificar la pérdida de masa asociada a una pérdida de energía. A partir de $\frac{\Delta E}{c^2} = \Delta m$, se puede estimar que el Sol pierde unas 4×10^6 toneladas de masa cada segundo[90].

Cuando un electrón se combina con su antipartícula, el positrón, de igual masa en reposo pero carga eléctrica opuesta, la carga desaparece, al igual que las masas, y aparecen fotones que conservan la energía y la cantidad de movimiento de las partículas originales. En esta transformación se basa la tomografía por emisión de positrones (PET). Un fotón en las proximidades de un núcleo atómico puede dar origen a un electrón y su antipartícula.

Así pues, **en física relativista la masa en reposo no se conserva**, a diferencia de lo que ocurre en física clásica. La masa puede desaparecer o aparecer. Solamente se conserva la energía.
Puesto que los experimentos sugieren que la masa inercial es igual a la masa gravitatoria, cuando una parte de la masa inercial desaparece, el peso de la masa remanente es inferior al peso de la masa inicial. Cabe preguntarse si la energía correspondiente a la masa faltante posee peso y si éste es el mismo que corresponde a dicha masa. Los experimentos sugieren que si: la proporcionalidad entre la masa inercial y la masa gravitacional se extiende a la masa relativista de una partícula y en

---

[89] R. Murray "*Nuclear energy: an introduction to the concepts, systems, and applications of nuclear processes*", Butterworth-Heinemann, Oxford, UK, 2008.

[90] Se puede hacer un cálculo aproximado de ΔE estimando la potencia por unidad de área que incide sobre nuestro planeta y suponiendo que la energía radiada se distribuye uniformemente en la superficie de una esfera con el sol en su centro y radio igual a la distancia entre la Tierra y el Sol.

general a toda forma de energía de un sistema de partículas en interacción, a la que pueda asociársele una masa equivalente (por ejemplo, energía potencial). Los primeros experimentos se realizaron entre 1910 y 1920, comparando los períodos de dos tipos de péndulo. Se experimentó con péndulos con masas involucrando núcleos atómicos radioactivos y péndulos con masas formadas por átomos cuyos núcleos son estables, sin hallar diferencias en los períodos que pudieran asignarse a diferencias en el peso de la energía[91].

Debido a la homogeneidad del espacio, las propiedades de un sistema aislado de partículas no cambian si el sistema sufre una traslación arbitraria (se mueve paralelamente a sí mismo una distancia cualquiera). Una consecuencia de esto es la conservación de una cantidad vectorial, la cantidad de movimiento total del sistema.
Si dos partículas de masas en reposo m_1 y m_2 se aproximan y sufren una colisión, dando origen a dos partículas de masas en reposo m_3 y m_4 , la conservación relativista de la cantidad de movimiento se formula así:

$$\vec{p}_1 + \vec{p}_2 = \frac{m_1}{\sqrt{1-\frac{v^2}{c^2}}} \cdot \vec{v}_1 + \frac{m_2}{\sqrt{1-\frac{v^2}{c^2}}} \cdot \vec{v}_2 = \frac{m_3}{\sqrt{1-\frac{v^2}{c^2}}} \cdot \vec{v}_3 + \frac{m_4}{\sqrt{1-\frac{v^2}{c^2}}} \cdot \vec{v}_4 = \vec{p}_3 + \vec{p}_4$$

Cuando las velocidades de las partículas son pequeñas respecto de c, se reduce a la fórmula de la mecánica newtoniana de la parte 2.2.
Las colisiones elásticas entre protones rápidos y protones prácticamente en reposo, observadas en cámaras de niebla, muestran que las velocidades y los ángulos de los dos protones después de la colisión concuerdan con lo que se desprende de las fórmulas relativistas para la conservación de la energía y de la cantidad de movimiento.

La isotropía del espacio implica que las propiedades de un sistema aislado de partículas no deben variar cuando se lo someta a una rotación arbitraria en torno a un eje cualquiera. Una consecuencia de esto es la conservación de una cantidad vectorial, el momento de la cantidad de movimiento o momento angular total del sistema, que depende de la velocidad de rotación y de la distribución de masas en torno al eje de giro. Si esta última se contrae hacia el eje, la velocidad angular debe aumentar para mantener constante el momento angular. Este proceso, como se verá en la sección 13.3.1, adquiere importancia durante la formación de una estrella de neutrones.

8.1.3 Transformación de la fuerza al cambiar de marco inercial. Acción y reacción

Consideremos una partícula que **en un instante dado** se mueve respecto de un marco K, paralelamente al eje x, con velocidad v sometida a una fuerza de componentes F_x, F_y y F_z respecto de los ejes de coordenadas cartesianas ortogonales. Supongamos que respecto del marco inercial de referencia K' la partícula se mueve con velocidad v' paralelamente al eje x', siendo x y x' paralelos, por acción de una fuerza de componentes $F_x^{'}, F_y^{'}$ y $F_z^{'}$ respecto de los ejes espaciales de K'.
Entonces es posible demostrar que se verifican las siguientes leyes de transformación[92]:

[91] W. Mc Crea "*Física relativista*", pp. 81-82, UTHEA, México, 1965.

[92] W. Mc Crea "*Física relativista*", pp. 30-32, UTHEA, México, 1965.

$$F_x^{'} = F_x \qquad \frac{F_y^{'}}{\sqrt{1-\frac{v'^2}{c^2}}} = \frac{F_y}{\sqrt{1-\frac{v^2}{c^2}}} \qquad \frac{F_z^{'}}{\sqrt{1-\frac{v'^2}{c^2}}} = \frac{F_z}{\sqrt{1-\frac{v^2}{c^2}}}$$

Supongamos que respecto de un marco K, en un cierto instante, dos partículas están sometidas a fuerzas iguales u opuestas, pero se mueven con diferente velocidad. Como consecuencia de esta diferencia de velocidad, respecto de otro marco K' estarán sometidas a fuerzas diferentes. Aplicando este resultado a las fuerzas de interacción entre dos partículas resulta que en general acción y reacción no serán iguales: la tercera ley de Newton no puede extenderse a la mecánica relativista.

Si se admite que la tercera ley se aplica (al menos como una muy buena aproximación) cuando las velocidades de movimiento respecto de un marco inercial son despreciables en comparación con la velocidad de la luz, entonces la tercera ley debería resultar aplicable en física relativista para un sistema de referencia en el caso límite en el que las dos partículas que interactúan poseen velocidad relativa nula.

Otra consecuencia interesante es que si respecto de un marco K la fuerza depende solamente de la posición de la partícula, respecto de otro marco K' la fuerza correspondiente depende en general de la posición y de la velocidad.

8.2 Electrodinámica, relatividad y cuantos

Hemos visto que si existe un límite superior finito para la propagación de las perturbaciones, no pueden existir cuerpos rígidos en sentido estricto, ni siquiera como una idealización teórica, puesto que en un cuerpo rígido todos los puntos deben comenzar a moverse en el mismo instante en el que se aplica una fuerza, no importando su distancia al punto de aplicación de la fuerza.

Esto tiene a su vez consecuencias sobre el concepto de partícula elemental, entendiendo por tal un cuerpo que participa en todos los movimientos posibles como una entidad, o sea, que carece de sentido hablar de sus partes. Su estado, desde el punto de vista de la cinemática, viene dado por su posición y por su velocidad como un todo.

Si una partícula elemental, tal como un electrón, poseyera dimensiones finitas, no podría deformarse sin dejar de ser elemental. Como la teoría de la relatividad elimina los cuerpos rígidos, **una partícula elemental debe necesariamente ser considerada en todos los casos estrictamente como un punto geométrico.** Es decir, no se trata de la posibilidad de despreciar las dimensiones de un cuerpo material para algunos problemas de movimiento, como ocurre en mecánica newtoniana, sino de la imposibilidad de describir la posición de una partícula elemental sin reducirla a un punto del espacio.

En 4.1 vimos que la auto-energía electroestática de una esfera de radio R_0 uniformemente cargada es $\frac{q^2}{R_0}$. Como un electrón se considera una partícula elemental, entonces para él $R_0 = 0$, por lo cual debe tener auto energía infinita. Como la masa en reposo se vincula con la energía de reposo E_0 de un cuerpo a través de la fórmula $m_0 = \frac{E_0}{c^2}$, y la energía debería incluir a la auto-energía electroestática, se obtendría un resultado físicamente absurdo: la masa en reposo del electrón debe ser infinita.

Puesto que la relación entre la masa y la energía se encuentra bien establecida, se concluye que la electrodinámica clásica no es aplicable a las partículas elementales,

partículas que desde la perspectiva de la física relativista deben ser consideradas como puntuales.
En particular resulta imposible discutir el problema del origen electromagnético de la masa del electrón en el marco de la electrodinámica clásica.
Si admitimos que la auto-energía electromagnética de un electrón es del orden de su energía en reposo $E_0 = m_0 \cdot c^2$, y si introducimos un radio r_0 tal que $m_0 \cdot c^2 \approx \frac{e^2}{r_0}$, obtenemos $r_0 \approx \frac{e^2}{m_0 \cdot c^2}$.
La distancia r_0 a la cual la auto-energía electroestática de un electrón se hace del mismo orden que la energía asociada con su masa de reposo se puede tomar como estimación de los límites de aplicabilidad de la electrodinámica clásica[93]: $r_0 \approx 10^{-15} m$
Cuando se tienen en cuenta los fenómenos cuánticos se obtiene una estimación de los límites de aplicabilidad de la electrodinámica clásica tres órdenes de magnitud mayor que la que obtuvimos a partir de la auto-energía electroestática, siendo h la constante de Planck: $r_{cuántica} \approx \frac{h}{m_0 \cdot c} \approx 10^{-12} m$
Por tanto parece que es la naturaleza cuántica del micro-mundo, y no la singularidad en la energía electromagnética para las partículas elementales, la que limita la validez de la electrodinámica clásica cuando entran en juego distancias muy pequeñas.

8.3 Partículas y corpúsculos: fotones, campos gravitatorios y relojes

A partir de las fórmulas para la cantidad de movimiento y la energía se desprende esta relación independiente del valor de la masa en reposo: $\vec{p} = \frac{E}{c^2} \cdot \vec{v}$
La cantidad de movimiento de una partícula puede permanecer constante si su masa en reposo tiende a cero y su velocidad se aproxima a la de la luz de forma tal que la **masa relativa** $m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}$ permanece constante e igual μ. En el límite la cantidad de movimiento es $\vec{p} = \mu \cdot \vec{c}$ y la energía $E = \mu \cdot c^2$ Entonces $p = \frac{E}{c}$.
Resulta conveniente reservar el término partícula para el caso en que la masa en reposo no es nula, y corpúsculo para abarcar tanto las partículas como los objetos a los que pueda asociarse una masa en reposo nula. Así pues, **la dinámica relativista permite la existencia de corpúsculos con masa en reposo nula moviéndose con la velocidad de la luz en el vacío**.
La teoría cuántica de la luz interpreta los fenómenos luminosos en términos de corpúsculos de este tipo: los fotones. La energía del fotón viene dada en función de su frecuencia f y de la constante de Planck h por la fórmula: $E = h \cdot f$

[93] Si la masa en reposo se interpreta mediante el modelo de Lorentz para un electrón acelerado considerado en 4.3, este parámetro coincide con el radio clásico del electrón introducido por Lorentz.

La masa relativa del fotón $\mu = \frac{E}{c^2} = \frac{h \cdot f}{c^2}$, su masa en reposo es nula y su cantidad de movimiento es $\vec{p} = \mu \cdot \vec{c} = \frac{h \cdot f}{c^2} \cdot \vec{c}$ o bien $p = \mu \cdot c = \frac{h \cdot f}{c}$.

Comparando **el cociente de frecuencias** de un mismo fotón respecto de dos marcos de referencia inerciales, por un lado obtenido (el cociente de frecuencias) de la óptica relativista a partir de los aspectos ondulatorios del fotón, y por el otro lado obtenido a partir de la dinámica relativista, se obtiene el mismo resultado[94]. Como consecuencia, la hipótesis de los fotones (cuantos de luz) resulta compatible con el principio de relatividad de Einstein.

Si suponemos que el fotón presenta una masa gravitatoria, que la masa relativa del fotón $\frac{h \cdot f}{c^2}$ es igual a su masa gravitatoria y que la energía $E = h \cdot f$ se debe interpretar como una energía cinética, entonces una aplicación algo ingenua[95] del teorema de la conservación de la energía suministra la siguiente igualdad entre las frecuencias de un mismo fotón cuando pasa por dos puntos P y P' con diferente potencial gravitatorio newtoniano $\varphi(P)$ y $\varphi(P')$:

$$h \cdot f(P) + \frac{h \cdot f(P)}{c^2} \cdot \varphi(P) = h \cdot f(P') + \frac{h \cdot f(P')}{c^2} \cdot \varphi(P')$$

En lo que sigue supondremos que el potencial se debe a una masa respecto a la cual los observadores situados en P y en P' se encuentran en reposo. Si se puede suponer simetría esférica, y si r_P es la distancia entre el centro de la masa gravitante y el punto P, entonces $\varphi(P) = -\frac{G \cdot M}{r_P}$. Eliminando la constante de Planck y reordenado se obtiene:

$$\frac{f(P')}{f(P)} = \frac{1 + \frac{1}{c^2} \cdot \varphi(P)}{1 + \frac{1}{c^2} \cdot \varphi(P')}$$

Si el potencial gravitatorio en el punto P' es mayor que el potencial gravitatorio en el punto P, entonces la frecuencia del fotón cuando pasa por P' es menor que su frecuencia cuando pasa por P: un observador situado en P' detecta un corrimiento hacia el rojo respecto de la frecuencia que mide un observador situado en P, o en forma equivalente, en P se detecta un corrimiento hacia el azul respecto de la frecuencia medida en P'.

Puesto que como se verá en el próximo capítulo, esta fórmula se puede justificar en el límite de campos débiles, resulta que $\frac{1}{c^2} \cdot \varphi$ debe ser muy pequeño respecto de la unidad. Entonces el **corrimiento relativo** se puede estimar mediante la fórmula:

$$\frac{f(P') - f(P)}{f(P)} \cong -\frac{(\varphi(P') - \varphi(P))}{c^2} = \frac{1}{c^2} \cdot G \cdot M \cdot \left(\frac{1}{r_{P'}} - \frac{1}{r_P} \right)$$

Esta expresión se aplica en astronomía para determinar el radio y la densidad de las estrellas.

---

[94] Esta comparación fue realizada por Kermack, Mc Crea y Whittaker en 1933. Ver W. Mc Crea "*Física relativista*", pp. 48-50, UTHEA, México, 1965.

[95] Pero que se puede justificar sobre la base del principio de equivalencia de Einstein.

En la Tierra se mide el corrimiento al rojo de una línea espectral de absorción conocida (línea de Fraunhofer) cuando el correspondiente fotón es emitido y luego absorbido desde la superficie de una estrella (por ejemplo, Sirio, que da un corrimiento relativo de -7×10^{-4}).

El potencial gravitatorio en la superficie de la estrella $\varphi(P) = -\frac{G \cdot M}{R}$ se puede estimar a partir de la ecuación para el corrimiento relativo, suponiendo que el punto P' (situado en la superficie de la Tierra) está tan alejado de la estrella que $\varphi(P') \approx 0$ y teniendo en cuenta que c = $2{,}99776 \times 10^{10}$ cm/s.

Cuando se conoce el potencial y se puede estimar la masa de la estrella ($M = 2\times 10^{33}$ gramos en el caso de Sirio), teniendo en cuenta que G = $6{,}67 \times 10^{-8}$ $cm^3/g \times s$, se puede calcular el radio R de la estrella (el cálculo da 2×10^8 cm para Sirio).

A esa masa y ese radio corresponden la densidad enorme del tipo de estrella conocida como enana blanca, en este caso de una densidad aproximada de 10^8 gramos/ cm^3.

Un átomo que emite una línea espectral de frecuencia f puede ser considerado como un reloj de período $T = \frac{1}{f}$.

Una consecuencia del efecto del campo gravitatorio sobre la frecuencia, y por tanto sobre el período del reloj atómico, es que un observador que examina la marcha de un reloj (es decir, de un átomo de un isótopo de un elemento químico) situado en un punto con menor potencial gravitatorio que el correspondiente al punto donde se halla el observador, constatará que ese reloj atrasa respecto de un reloj formado por un átomo del mismo isótopo, cuando este último está colocado junto al observador.

Cabe esperar el mismo efecto del campo gravitatorio sobre cualquier proceso periódico.

8.4 Invariancia de la fase de una onda

La fase de una onda permite caracterizar eventos tales como el desvanecimiento de los campos eléctrico y magnético observados respecto de un marco K en un instante t y en un punto de coordenadas x,y,z. Respecto de un nuevo marco K' ese evento se producirá en un instante t' y en un punto de coordenadas x', y', z', pero el evento en sí mismo, es decir, la anulación de los campos en el caso del ejemplo que consideramos, no habrá cambiado. De lo contrario el marco K respecto del cual los campos se desvanecen se podría distinguir de todos los demás marcos inerciales a partir de la observación de un fenómeno físico, violando entonces el principio de relatividad de Einstein.

8.4.1 Efecto Doppler y aberración relativista de la luz

Consideremos entonces la fase de una onda armónica plana representada respecto del marco K en un punto de coordenadas espaciales x, y, z en el instante t:

$$\Phi(t, x, y, z) = x \cdot k_x + y \cdot k_y + z \cdot k_z - \omega \cdot t$$

Respecto de otro marco K' la fase verifica:

$$x \cdot k_x + y \cdot k_y + z \cdot k_z - \omega \cdot t = x' \cdot k'_x + y' \cdot k'_y + z' \cdot k'_z - \omega' \cdot t'$$

Sustituyendo x', y', z' y t' en función de x, y, z, y t de acuerdo con las fórmulas de transformación de Lorentz, y comparando coeficientes entre el miembro de la izquierda y el miembro de la derecha de la igualdad, se obtienen las fórmulas de transformación para las componentes k_x, k_y, k_z del número de ondas y para la frecuencia angular ω de oscilación: $k_x = \dfrac{k'_x + \dfrac{\omega' \cdot V}{c^2}}{\sqrt{1-\dfrac{V^2}{c^2}}}$ $\quad k_y = k'_y \quad k_z = k'_z \quad \omega = \dfrac{\omega' + k'_x \cdot V}{\sqrt{1-\dfrac{V^2}{c^2}}}$

Se advierte que la componente del vector de ondas paralela a la velocidad relativa de los marcos se transforma como el tiempo y la frecuencia como la coordenada espacial correspondiente

Supongamos una fuente de ondas fija en K', de frecuencia ω', y admitamos que θ' es el ángulo entre la velocidad relativa V de K' respecto de K y el segmento de recta que une un receptor fijo a K con la fuente. En ese caso: $k'_x = k' \cdot \cos\theta' = \dfrac{\omega'}{c} \cdot \cos\theta'$

Sustituyendo este resultado en la fórmula de transformación para la frecuencia se obtiene:

$$\omega = \frac{\omega' \cdot \left(1 + \dfrac{V}{c} \cdot \cos\theta'\right)}{\sqrt{1-\dfrac{V^2}{c^2}}}$$

Si la fuente se mueve hacia el observador, $\theta' = \pi$ y entonces se obtiene una fórmula utilizada para determinar las velocidades radiales de las estrellas:

$$\omega = \frac{\omega' \cdot \left(1 - \dfrac{V}{c}\right)}{\sqrt{1-\dfrac{V^2}{c^2}}} = \omega' \cdot \sqrt{\frac{1-\dfrac{V}{c}}{1+\dfrac{V}{c}}}$$

En el caso no relativista se desprecia $\dfrac{V^2}{c^2}$ respecto de la unidad y $\omega = \omega' \cdot \left(1 - \dfrac{V}{c}\right)$

La corrección relativista viene dada por el factor $\dfrac{1}{\sqrt{1-\dfrac{V^2}{c^2}}}$

Si la velocidad de la fuente es perpendicular al observador, $\theta' = \dfrac{\pi}{2}$ y en este caso:

$$\omega = \frac{\omega'}{\sqrt{1-\dfrac{V^2}{c^2}}}$$

En el caso no relativista las frecuencias son iguales. En el caso relativista las frecuencias ya no son iguales. Hay un efecto, pero es de segundo orden respecto de $\dfrac{V}{c}$. Este efecto

fue observado por Ives[96] en la radiación proveniente de iones en movimiento rápido como para que $\frac{V}{c}$ resulte lo suficientemente próximo a 1 como para poder detectar el desplazamiento de frecuencia por métodos espectroscópicos.

Consideremos ahora un rayo de luz emitido en el marco K' formando un ángulo θ' con el eje x'. Igual que antes, suponemos que el marco K' se mueve respecto del marco K con velocidad V manteniéndose paralelos los ejes x y x'. Entonces el rayo forma un ángulo θ' también respecto de eje x.

La magnitud del número de ondas cambia dependiendo del marco de referencia respecto del cual se lo considere, puesto que $k = \frac{\omega}{c}$ y $k' = \frac{\omega'}{c}$, siendo $\frac{\omega'}{\omega} = \frac{\sqrt{1-\frac{V^2}{c^2}}}{1+\frac{V}{c}\cdot\cos\theta'}$

De la igualdad $k_y = k'_y$ se desprende que $k \cdot sen\theta = k' \cdot sen\theta'$, es decir que el ángulo θ del rayo de luz respecto del eje x va a ser diferente de θ': es el fundamento de la corrección relativista al efecto de aberración de la luz ya estudiado en óptica clásica y aplicado en astronomía.

Busquemos entonces una relación entre θ y θ'.

Comencemos con $sen\theta = \frac{k'}{k}\cdot sen\theta' = \frac{\omega'}{\omega}\cdot sen\theta'$ Teniendo en cuenta la fórmula para $\frac{\omega'}{\omega}$ introducida más arriba, resulta finalmente: $sen\theta = \frac{\sqrt{1-\frac{V^2}{c^2}}}{1+\frac{V}{c}\cdot\cos\theta'}\cdot sen\theta'$

Esta última ecuación resulta de importancia al considerar el próximo tema: cómo aparecen cuando son vistos por un ojo o fotografiados por una cámara los objetos en rápido movimiento respecto de la cámara o del ojo.

8.4.2 La apariencia visual de objetos que se mueven con velocidades cercanas a la de la luz en el vacío

Desde los trabajos de Einstein de 1905 hasta los trabajos de James Terrell y Roger Penrose de 1959, la gente pensaba que durante una observación visual un objeto en movimiento se vería contraído en la dirección de su velocidad V por un factor $\sqrt{1-\frac{V^2}{c^2}}$.

Como consecuencia el efecto solo sería aparente a velocidades lo bastante próximas a la de la luz y en esas condiciones se creía que una esfera parecería un elipsoide achatado en su dirección de movimiento.

[96] H. Ives y G. Stilwell, "An experimental study of the rate of a moving atomic clock", Journal of the Optical Society of America, 28, 1938, pp.215-226.

No obstante, Penrose y Terrell (y este último en particular) demostraron que el efecto del movimiento aparece como una rotación de los cuerpos respecto de su apariencia en reposo[97].

Cuando vemos o tomamos una instantánea con una cámara fotográfica los cuantos de luz que llegan en forma simultánea a nuestra retina o a la placa fotográfica, no han sido emitidos simultáneamente desde el objeto: los puntos más alejados del ojo o de la cámara han emitido su parte de la imagen antes de que los puntos más próximos hayan emitido la suya.
Como un objeto en movimiento respecto de un observador se encuentra en puntos diferentes en instantes diferentes cuando las distintas partes han emitido la parte de la imagen que les corresponde, esto tiene como consecuencia una imagen distorsionada del objeto que está siendo observado visualmente[98].
Lo que Terrell descubrió es que para objetos lo bastante alejados como para que los rayos luminosos que de ellos provienen se puedan considerar paralelos, la contracción de Lorentz-FitzGerald cancela parte de la distorsión de modo tal que los objetos de forma tal que estos últimos no aparecen distorsionados, sino que **aparecen rotados**.
Consideremos un conjunto de pulsos de luz que salen de puntos diferentes de un mismo objeto y viajan paralelos entre sí en dirección a un observador. Supongamos que en el marco K respecto del cual tanto el observador como el objeto están en reposo, los

---

[97] Una revisión del problema de la apariencia visual de los cuerpos en movimiento puede verse en V. Weisskopf "*La física en el siglo XX*", Alianza editorial, Madrid, 1990, pp. 226-235.

[98] Supongamos que una varilla de longitud l lejana respecto de un observador, se mueve perpendicularmente a su eje longitudinal con velocidad uniforme V y pasa justo por encima del ojo de un observador.
Si la velocidad de la luz fuera infinita, el observador vería una imagen de punto o pequeño círculo.
Pero como es finita, la luz emitida por un punto del extremo distal de la varilla (alejado del observador) tarda un tiempo $\frac{l}{c}$ en alcanzar el extremo proximal. Dos pulsos de luz, uno proveniente del extremo distal y otro del proximal de la varilla, que alcanzan simultáneamente el ojo del observador, no fueron emitidos simultáneamente. El que viene del extremo distal fue emitido un tiempo $\frac{l}{c}$ antes que el otro. Pero en ese momento la varilla no se encontraba justo encima del observador, sino en un punto situado a una distancia $\frac{l}{c}.V$ de esa posición. Entonces el pulso llega al ojo proveniente de ese otro punto. Lo mismo ocurre con los demás puntos de la varilla: el pulso proveniente de un punto situado a una distancia x del extremo proximal y que llega al ojo simultáneamente con los demás, fue emitido un tiempo $\frac{x}{c}$ antes y desde un punto situado a una distancia $\frac{x}{c}.V$ Entonces la imagen de la varilla en movimiento no es un punto, sino un segmento de recta de longitud $\frac{l}{c}.V$ paralelo a la dirección de movimiento y situado por atrás del punto que corresponde al extremo proximal de la varilla.
La varilla aparece entonces **rotada hacia atrás** un ángulo φ respecto de su posición "real", tal que $sen\varphi = \frac{V}{c}$

pulsos alcanzan en forma simultánea un plano ortogonal a su dirección de movimiento, que denominaremos plano del observador, y forman allí una imagen del objeto.
Si el objeto comienza a moverse con velocidad V respecto del observador, solidario con un marco K', los rayos provenientes del objeto forman un ángulo θ' respecto del eje x' de K', pero forman un ángulo diferente θ respecto de K. La fórmula de la aberración de la luz deducida en la sección 8.4.1 conecta estos ángulos entre sí.

Según la física pre-relativista, debido a que asume la validez de la ley de adición de velocidades de Galileo, los pulsos que antes llegaban simultáneamente al plano del observador cuando éste y el objeto observado se encontraban en reposo respecto del mismo marco K, ahora no lo hacen y en consecuencia la imagen se distorsiona.
Pero según la física relativista, cuando el objeto se mueve respecto del observador, los pulsos de luz también llegarían simultáneamente al plano del observador y mantendrían las mismas distancias relativas que guardaban al llegar a ese plano cuando el objeto y el observador se encontraban en reposo relativo: en consecuencia la imagen del objeto no se distorsiona con el movimiento.
Pero entonces lo único que ha variado es el ángulo de los rayos paralelos provenientes del objeto: pasa de θ' cuando el objeto no se mueve a θ cuando lo hace respecto del observador. Por tanto la imagen que se forma en el plano del observador aparece girada un ángulo $\theta - \theta'$.

8.5 Cuadrivectores, dinámica relativista y campo electromagnético

Asociados con los eventos en el espacio de cuatro dimensiones de Minkowski, se pueden introducir vectores tetra-dimensionales denominados **cuadrivectores**.
Los cuadrivectores resultan muy útiles para formular la dinámica relativista en el continuo tiempo-espacio, porque en forma natural sustituyen la descripción de las propiedades de los sistemas definidas en el espacio físico tridimensional por otras que las generalizan en el continuo tiempo-espacio de cuatro dimensiones.
Respecto de un marco K los cuadrivectores poseen, como su nombre lo indica, cuatro componentes: una de ellas se denomina **temporaloide** porque se transforma como el tiempo al cambiar de marco de referencia, y las otras tres **espacialoides**, porque se transforman como el vector de posición.
Si $\tilde{A}$ es un cuadrivector, respecto de K se representa así: $\left(\tilde{A}\right)_K = \left(a_t, \vec{a}_r\right)$
En esta representación a_t es la componente temporaloide y $\vec{a}_r$ representa las tres componentes espacialoides del cuadrivector. Respecto de otro marco: $\left(\tilde{A}\right)_{K'} = \left(a'_t, \vec{a}'_r\right)$
Generalizando las fórmulas presentadas a comienzo del capítulo 7 para la transformación de las coordenadas de tiempo-espacio $\left(t, \vec{r}\right)$ de un evento, las fórmulas de Lorentz pueden escribir, para la transformación de las componentes de cualquier cuadrivector, de la siguiente forma:

$$a'_t = \gamma \cdot \left(a_t - \frac{\vec{a}_r \bullet \vec{V}}{c} \right) \qquad \vec{a}'_r = \vec{a}_r + (\gamma - 1) \cdot \left(\frac{\vec{a}_r \bullet \vec{V}}{V^2} \right) \cdot \vec{V} - \gamma \cdot a_t \cdot \frac{1}{c} \vec{V}$$

A partir de estas fórmulas se demuestra que el escalar $a_t^2 - \vec{a}_r \bullet \vec{a}_r$ permanece invariante al cambiar de un marco de referencia inercial a otro.

Un primer ejemplo es el cuadrivector de posición témporo-espacial de un evento, que respecto de un marco de referencia K viene dado por: $\left(\tilde{R}\right)_K = \left(c \cdot t, \vec{r}\right)$

Utilizando el tiempo propio τ como parámetro se define una cuadrivelocidad:

$$\left(\tilde{U}\right)_K = \left(\frac{d\tilde{R}}{d\tau} \right)_K = \left(c \cdot \frac{dt}{d\tau}, \frac{dt}{d\tau} \cdot \frac{d\vec{r}}{dt} \right)$$

En esta representación, si v es la velocidad de la partícula respecto del marco K, $\frac{dt}{d\tau} = \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$ (Se supone que esta fórmula se verifica también en el caso de una partícula acelerada.) Por tanto, la componente temporaloide de la cuadrivelocidad es $\frac{dx_0}{d\tau} = \gamma \cdot c$ mientras que las componentes espacialoides verifican $\frac{dx_i}{d\tau} = \gamma \cdot \frac{dx_i}{dt}$ siendo $i = 1,2,3$. Así pues: $\left(\tilde{U}\right)_K = \left(\gamma \cdot c, \gamma \cdot \vec{v}\right)$

Entonces el producto escalar $\left(\tilde{U}\right)_K \bullet \left(\tilde{U}\right)_K = \gamma^2 \cdot c^2 - \gamma^2 \cdot v^2 = c^2$ es invariante (no depende del marco de referencia K).

Otro ejemplo de cuadrivector es $\tilde{P}$, el cuadrivector de energía-cantidad de movimiento lineal para una partícula: $\left(\tilde{P}\right)_K = \left(\frac{E}{c}, \vec{p} \right)$ donde $E = m_0 \cdot \gamma \cdot c^2 \qquad \vec{p} = m_0 \cdot \gamma \cdot \vec{v}$

Puesto que la masa propia se considera es invariante bajo las transformaciones de Lorentz, $\left(\tilde{P}\right)_K \bullet \left(\tilde{P}\right)_K = \left(\frac{E}{c} \right)^2 - \vec{p} \bullet \vec{p} = m_0^2 \cdot c^2$ es independiente del marco de referencia.

De sus definiciones se desprende que la cuadrivelocidad y el cuadrivector de energía-cantidad de movimiento se relacionan con la masa en reposo de la partícula así:

$$\left(\tilde{P}\right)_K = m_0 \cdot \left(\tilde{U}\right)_K$$

Esta fórmula generaliza la relación $\vec{p} = m \cdot \vec{v}$ entre la velocidad y la cantidad de movimiento de la mecánica newtoniana.

En mecánica relativista se puede introducir un cuadrivector de potencia-fuerza derivando el cuadrivector de energía-cantidad de movimiento de una partícula respecto del tiempo propio de esa partícula e introduciendo las conexiónes relativistas entre la variación en la cantidad de movimiento y la fuerza, y la variación en la energía y potencia, estudiadas en 8.1: $\left(\tilde{F}\right)_K = \frac{d}{d\tau} \left(\tilde{P}\right)_K$

La componente temporaloide del nuevo cuadrivector es $\frac{1}{c}\cdot\frac{dE}{d\tau}=\frac{\gamma}{c}\cdot\frac{dE}{dt}=\frac{\gamma}{c}\left(\vec{F}\cdot\vec{v}\right)$ mientras que las componentes espacialoides forman el vector $\frac{d\vec{p}}{d\tau}=\gamma\cdot\frac{d\vec{p}}{dt}=\gamma\cdot\vec{F}$

Esta formulación modificada de la versión relativista de la segunda ley de Newton, empleando cuadrivectores y el tiempo propio asociado a la partícula, se considera superior a la presentada en 8.1, puesto que como el tiempo propio es un invariante, mantiene la misma forma respecto de todos los marcos inerciales: lo único que cambia es la descomposición de los cuadrivectores. Esta descomposición depende en cada caso del marco de referencia de tiempo-espacio.

Las fórmulas de transformación de las componentes temporaoide y espacialoide de un cuadrivector, aplicadas al cuadrivector de potencia fuerza $\left(\tilde{F}\right)_K$ determinan cómo se transforma la fuerza $\vec{F}$ al cambiar de marco de referencia inercial. Se obtiene así un resultado que generaliza el presentado en 8.1.3.

Mientras que la fuerza que el campo electromagnético ejerce sobre una carga en movimiento (la fuerza de Lorentz) se transforma de esta forma y por lo tanto es compatible con la teoría restringida de la relatividad, la fuerza de interacción gravitatoria dada por la ley de Newton no lo hace. Por este motivo esta fuerza de atracción gravitatoria no se puede incorporar en un cuadrivector de potencia-fuerza.

En la teoría de Maxwell las cargas y sus movimientos son las fuentes del campo electromagnético. Se supone que la carga eléctrica es un invariante relativista, al igual que el número de partículas cargadas que componen un cierto sistema (cuando no se producen reacciones de formación o de aniquilación de pares).

Entonces, introduciendo un marco auxiliar K_0 localmente en co-movimiento con un grupo de cargas, es posible deducir expresiones para la densidad de carga $\rho_q(t,\vec{r})$ y la densidad de corriente eléctrica $\vec{J}_q(t,\vec{r})$ respecto de un marco de referencia K y la forma en que se transforman (en $\rho'_q(t',\vec{r}\,')$ y $\vec{J}'_q(t',\vec{r}\,')$) al pasar a describir los fenómenos respecto de un nuevo marco K'[99]. Se encuentra así que la densidad de carga se transforma como el tiempo y las componentes de la densidad de corriente se transforman como las componentes del vector de posición. Se puede formar un cuadrivector $\left(\rho_q,\vec{J}_q\right)_K$.

En electrodinámica se demuestra que el campo eléctrico $\vec{E}(t,\vec{r})$ y el campo magnético $\vec{H}(t,\vec{r})$ pueden expresarse en términos de un potencial escalar $\phi(t,\vec{r})$ y un potencial vectorial $\vec{A}(t,\vec{r})$ de componentes cartesianas $A_x(t,\vec{r})$ $A_y(t,\vec{r})$ $A_z(t,\vec{r})$:

$$\vec{E}(t,\vec{r})=-\nabla\phi-\frac{1}{c}\frac{\partial\vec{A}}{\partial t} \qquad \vec{H}(t,\vec{r})=\nabla\times\vec{A}$$

[99] Sobre las definiciones de densidad de carga y de corriente en física relativista, así como el resto de los temas básicos de la formulación covariante del electromagnetismo, ver por ejemplo W. Mc Crea, "*Física relativista*", capítulo 8, UTHEA, México, 1965.

Como es usual en análisis vectorial, el producto vectorial $\nabla \times \vec{A}$ del operador ∇ por el campo $\vec{A}$ representa al rotor del campo.

A partir de los potenciales escalar y vectorial se puede definir un cuadrivector $\left(\tilde{A}\right)_K = \left(\phi, \vec{A}\right)_K$.

Introduciendo la conexión $\frac{\partial A_x(t,\vec{r})}{\partial x} + \frac{\partial A_y(t,\vec{r})}{\partial y} + \frac{\partial A_z(t,\vec{r})}{\partial z} + \frac{\partial \phi}{\partial t} = 0$, a partir de las ecuaciones de Maxwell se deduce que la componente temporaloide $\phi(t,\vec{r})$ verifica una ecuación de ondas clásica cuyo término fuente es la densidad de carga, y que la componente espacialoide $\vec{A}(t,\vec{r})$ verifica también una ecuación de ondas con la densidad de corriente como término fuente del campo.

A partir del análisis del cuadrivector formado con los potenciales del campo y de las relaciones entre los potenciales y los campos se demuestra que los campos se transforman según las siguientes fórmulas al cambiar de marco de referencia:

$$\vec{E}' = \gamma \cdot \vec{E} + (1-\gamma) \cdot \frac{\vec{E} \bullet \vec{V}}{V^2} \cdot \vec{V} + \gamma \cdot \frac{\vec{V} \times \vec{H}}{c}$$

$$\vec{H}' = \gamma \cdot \vec{H} + (1-\gamma) \cdot \frac{\vec{H} \bullet \vec{V}}{V^2} \cdot \vec{V} - \gamma \cdot \frac{\vec{V} \times \vec{E}}{c}$$

Estas fórmulas se pueden comparar con la que corresponde a la transformación de la componente espacialoide de un cuadrivector y resulta que son diferentes: desde el punto de vista relativista, $\vec{E}$ y $\vec{H}$ no son propiamente vectores.

Un análisis más profundo arroja como resultado que $\vec{E}$ y $\vec{H}$ deben considerarse siempre en conjunto, como integrantes de un objeto más complejo, el tensor del campo electromagnético. En particular, las fórmulas de transformación muestran que la presencia o la ausencia de $\vec{E}$ o de $\vec{H}$ tomados por separado depende del marco de referencia respecto del cual se describan esos campos.

Resulta interesante que si el campo eléctrico se anula en un marco K, es decir $\vec{E} = \vec{0}$, entonces en cualquier otro marco K' se tiene:

$$\vec{E}' = \gamma \cdot \frac{\vec{V} \times \vec{H}}{c} \qquad \vec{H}' = \gamma \cdot \vec{H} + (1-\gamma) \cdot \frac{\vec{H} \bullet \vec{V}}{V^2} \cdot \vec{V}$$

Como el triple producto $\vec{a} \bullet \left(\vec{b} \times \vec{c}\right)$ se anula si dos factores son iguales, resulta que basta que respecto de un marco K el campo eléctrico sea nulo, para que respecto de otro marco K' cualquiera, los campos eléctrico y magnético **ser perpendiculares entre sí**: $\vec{E}' \bullet \vec{H}' = 0$

8.6 La fórmula para el tiempo propio de los relojes acelerados[100]

En la formulación de la dinámica relativista utilizando cuadrivectores y el tiempo propio como parámetro, subsiste el problema, señalado en 7.3, que se plantea al utilizar la misma fórmula para el tiempo propio basada en la relación entre marcos inerciales para conectar un marco acelerado (el marco fijo a la partícula) con un marco inercial (respecto del cual la partícula posee la velocidad $\vec{v}$).

Se suele argumentar que en cada instante medido respecto del marco inercial K se puede identificar un marco de referencia inercial que posee la misma velocidad $\vec{v}$ que la partícula en ese instante. Como esta velocidad puede variar con el paso del tiempo en el marco K, entonces el marco inercial en el que la partícula se encuentra instantáneamente en reposo debe variar: se tiene una familia de marcos inerciales tangentes, por así decirlo.

Entre el incremento de tiempo propio $\Delta\tau$ que separa dos eventos que ocurren en el mismo lugar respecto de uno de estos marcos tangentes y el incremento de tiempo Δt medido en K que separa esos mismos dos eventos, se relacionan a través de la fórmula:

$$\frac{\Delta t}{\Delta\tau} = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\,.$$

Luego se hace tender $\Delta\tau$ a cero y se interpreta el resultado obtenido $\frac{dt}{d\tau} = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$, como la relación entre un tiempo propio asignable a la partícula y el tiempo medido en de un marco inercial respecto del cual se describe el movimiento.

Como se dijo en 7.3, esto debe ser introducido como una suposición adicional en el marco de la teoría restringida de la relatividad. Si se la admite, es necesario tener en cuenta que una de sus consecuencias es, al menos, problemática.

Puesto que $\tau_2 - \tau_1 = \int_{t_1}^{t_2} \sqrt{1-\frac{v^2}{c^2}} \cdot dt$, si la velocidad está acotada por un valor inferior a c, al tender a infinito $t_2 - t_1$, tiende también a infinito $\tau_2 - \tau_1$.

Pero si la velocidad, siempre estando acotada por c, se incrementara lo bastante rápido, la integral $\int_{t_0}^{\infty} \sqrt{1-\frac{v^2}{c^2}} \cdot dt$ podría ser convergente. Si así fuera, parte de la evolución del observador acelerado (cronometrada por su tiempo propio τ) debería ocurrir fuera del universo accesible al observador inercial (cronometrada por el tiempo t). A este último "se le acabaría el tiempo" ($t \to \infty$) mientras que el tiempo propio del observador acelerado se mantendría finito. En ese caso el observador acelerado vería desaparecer al

[100] Esta parte se basa en los artículos de J. Ferrari y R. Suárez-Ántola "About the slowing down of accelerated clocks", Lettere al Nuovo Cimento, vol. 44, Nº 8, 1985, pp. 599-600 y de R. Suárez Ántola, "Acerca del enlentecimiento de los relojes acelerados II: tiempo propio relativista y unus mundus físico", Rev. SCP, Tercera Época, 23(2): 199-226, 2018. DOI: https://doi.org/10.32480/rscp.2018-23-2.199-226

observador inercial. Veamos ahora bajo qué condiciones esto podría ocurrir, si se asume la validez irrestricta de la fórmula integral para el tiempo propio.

Considerando, para simplificar, el movimiento de una partícula en una dimensión espacial, a partir de las ecuaciones $\gamma = \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}}$ y $p = m_0 \cdot \gamma \cdot v$ se deduce

$$\int_{t_0}^{\infty} \sqrt{1 - \frac{v^2}{c^2}} \cdot dt = \int_{t_0}^{\infty} \left(1 + \frac{p^2(t)}{m_0^2 \cdot c^2}\right)^{-\frac{1}{2}} \cdot dt \qquad (1)$$

Para un movimiento en dirección de la coordenada cartesiana x la generalización relativista de la segunda ley de Newton es $\frac{dp(t)}{dt} = F(t)$ ($p(t)$ es la cantidad de movimiento y $F(t)$ la fuerza). Entonces $p(t) = p(t_0) + \int_{t_0}^{t} F(t').dt'$. Sustituyendo en el segundo miembro de (1) se obtiene:

$$\tau(t) - \tau(t_0) = \int_{t_0}^{t} \sqrt{1 - \frac{v^2(t')}{c^2}} \cdot dt' = \int_{t_0}^{t} \frac{1}{\sqrt{1 + \frac{\left(p(t_0) + \int_{t_0}^{t'} F(t') \cdot dt'\right)^2}{m_0{}^2 \cdot c^2}}} \cdot dt' \qquad (1\text{ a})$$

A partir de esta ecuación es posible demostrar que $\int_{t_0}^{\infty} \sqrt{1 - \frac{v^2}{c^2}} \cdot dt$ diverge *siempre y cuando* sea divergente $\int_{t_0}^{\infty} \left(\int_{t_0}^{t} F(t').dt'\right)^{-1} \cdot dt$ (*Condición de divergencia para fuerza dada en función del tiempo*).

Si, por ejemplo, $F(t) = \ln t$ esta condición se cumple, mientras que si $F(t) = t^p$ con $p > 0$ ya no se cumple y entonces $\int_{t_0}^{\infty} \sqrt{1 - \frac{v^2}{c^2}} \cdot dt$ converge.

Si el movimiento se produce en un campo de fuerzas estacionario $F(x)$, de la ley del movimiento se obtiene, luego de varias transformaciones:

$$\tau(x) - \tau(x_0) = \frac{1}{c} \cdot \int_{x_0}^{x} \left[\left(\sqrt{1 + \frac{p^2(x_0)}{m_0{}^2 \cdot c^2}} + m_0{}^{-1} \cdot c^{-2} \cdot \int_{x_0}^{x} F(x'') \cdot dx''\right)^2 - 1\right]^{-\frac{1}{2}} \cdot dx'$$

(2)

Esta integral diverge *siempre y cuando* la integral $\int_{x_0}^{\infty}\left(\int_{x_0}^{x} F(x')\cdot dx'\right)^{-1}\cdot dx$ sea divergente (*Condición de divergencia para fuerza dada en función de la coordenada espacial*).
Lo es, por ejemplo, para $F(x)=\ln x$, pero la integral converge para $F(x)=x^p$ con $p>0$.

Entonces, si admitimos que $\int_{t_0}^{\infty}\sqrt{1-\frac{v^2}{c^2}}\cdot dt$ no puede ser convergente (con el fin de evitar la paradoja que constituye la desaparición del observador inercial del universo del observador acelerado), solamente deberíamos encontrar en la naturaleza campos de fuerza medidos respecto del observador inercial y actuando sobre el observador acelerado para los cuales la integral sea divergente.
La alternativa para evitar la paradoja es que $\tau_2-\tau_1=\int_{t_1}^{t_2}\sqrt{1-\frac{v^2}{c^2}}\cdot dt$ constituye solo una buena aproximación a la verdadera ecuación de transformación, aproximación cuyos límites de validez no pueden ser establecidos en el marco de la teoría restringida de la relatividad.

Cuando se profundiza el análisis en el marco de la teoría restringida, nada especial (además de la tasa de crecimiento hacia el infinito) parece caracterizar el comportamiento de la componente tangencial de los campos de fuerza a lo largo de la trayectoria del reloj acelerado, en relación con la convergencia o divergencia de la integral que da el tiempo propio como función del tiempo inercial.
No obstante, visto desde el punto de vista de la aceleración propia del reloj [101], aparece una diferencia físicamente significativa entre los dos casos lógicamente posibles.
Uno está integrado por las historias de aceleración propia tangenciales que dan un tiempo propio finito para un tiempo inercial infinito.
El otro está integrado por las historias de aceleraciones propias tangenciales que dan un tiempo propio infinito para un tiempo inercial infinito.
En el primer caso aparece una singularidad en la aceleración propia para un valor finito de tiempo propio. En el segundo caso esa singularidad no aparece.

8.7 Algunos resultados relacionados con la mecánica de fluidos y la termodinámica relativistas

[101] R. Suárez Ántola, "Acerca del enlentecimiento de los relojes acelerados II: tiempo propio relativista y unus mundus físico", Rev. SCP, Tercera Época, 23(2): 199-226, 2018.
DOI: https://doi.org/10.32480/rscp.2018-23-2.199-226

El estudio de la mecánica de fluidos relativista presenta considerable interés científico, sobre todo en relación con ciertos problemas astrofísicos y cosmológicos.
Aunque ciertamente en menor medida, cabe esperar que durante este siglo presente un interés tecnológico creciente.
Cuando las velocidades del movimiento microscópico de las partículas del fluido no son despreciables respecto de la velocidad de la luz, las ecuaciones del movimiento deben tener en cuenta efectos relativistas aun cuando las velocidades macroscópicas del medio sean pequeñas en comparación con la de la luz.
A diferencia de lo ocurrido con las variables de la mecánica de fluidos no relacionadas con la termodinámica, las leyes de transformación de las variables termodinámicas han suscitado numerosas discusiones y desacuerdos que se iniciaron hace 100 años y se han extendido hasta el presente. Aún hoy la pregunta acerca de si una termodinámica relativista tiene sentido como disciplina física, y si lo tiene, cómo debe ser formulada, no posee una respuesta aceptada en forma unánime por la comunidad científica.
En la medida en que la termodinámica y la mecánica de fluidos se puedan aplicar en estudios sobre la estructura, el origen y la evolución del Universo, también presenta interés desde una perspectiva filosófica.

Comencemos con la presión, en el caso de un fluido ideal. Supongamos que el fluido se encuentra en reposo respecto de un marco de referencia inercial K' cuyo eje x' se mueve con velocidad v, paralelamente al eje x de otro marco de referencia inercial K. La presión respecto de K' se define como es usual y adopta un valor P' que en cada punto toma el mismo valor con independencia de la orientación del elemento de superficie utilizado para definirla. En particular, representando la fuerza de compresión por unidad de área ortogonal a un eje mediante un subíndice que representa a dicho eje, se tiene: $P'_x = P'_y = P'_z$
La fuerza paralela al eje x' debida a la presión ejercida sobre un elemento de área $\Delta y' \cdot \Delta z'$ viene dada por $F'_x = P'_x \cdot \Delta y' \cdot \Delta z'$. Las fuerzas ejercidas sobre los otros dos elementos de área son $F'_y = P'_y \cdot \Delta x' \cdot \Delta z'$ y $F'_z = P'_x \cdot \Delta x' \cdot \Delta y'$.
En lo que sigue supondremos que los elementos de área son físicamente infinitesimales. En ese caso el elemento de volumen $\Delta V' = \Delta x' \cdot \Delta y' \cdot \Delta z'$ también lo será.
Respecto del marco K tenemos las tres fuerzas debidas a la presión:

$$F_x = P_x \cdot \Delta y \cdot \Delta z \qquad F_y = P_y \cdot \Delta x \cdot \Delta z \qquad F_z = P_z \cdot \Delta x \cdot \Delta y$$

Aplicando las leyes de transformación mostradas en 8.1.3, teniendo en cuenta que como el fluido se encuentra en reposo respecto de K' se debe poner $v' = 0$, resulta:

$$F_x = F'_x \qquad F_y = \sqrt{1-\frac{v^2}{c^2}} \cdot F_y' \qquad F_z = \sqrt{1-\frac{v^2}{c^2}} \cdot F_z'$$

Por otro lado, teniendo en cuenta la contracción de Fitz-Gerald que aparece en dirección del eje x, se tienen las leyes de transformación de los elementos de área:

$$\Delta y \cdot \Delta z = \Delta y' \cdot \Delta z' \qquad \Delta x \cdot \Delta z = \sqrt{1-\frac{v^2}{c^2}} \cdot \Delta x' \cdot \Delta z' \qquad \Delta x \cdot \Delta y = \sqrt{1-\frac{v^2}{c^2}} \cdot \Delta x' \cdot \Delta y'$$

De estas leyes de transformación y de las fórmulas para las componentes de las fuerzas en función de las presiones se desprende que $P'_x = P_x \quad P'_y = P_y \quad P'_z = P_z$.
Pero como $P'_x = P'_y = P'_z = P'$, resulta que $P_x = P_y = P_z = P$ y por tanto $P' = P$: la presión es un invariante relativista, es decir, un verdadero escalar.

A su vez los elementos de volumen físicamente infinitesimales ΔV y $\Delta V'$ se encuentran conectados por la relación $\Delta V = \sqrt{1-\frac{v^2}{c^2}} \cdot \Delta V'$

Todo lo anterior desde un punto de vista de mecánica de medios continuos. Desde el punto de vista microscópico propio de la teoría cinética molecular, la presión se interpreta como la velocidad de transferencia de cantidad de movimiento por unidad de área de una pared. Aplicando las leyes de transformación relativistas se deduce nuevamente que la presión es un invariante[102].

Si el volumen del sistema material es finito, y si dos observadores, situados en marcos de referencia diferentes, definen el volumen a partir de la región que para cada uno de ellos se encuentra ocupada simultáneamente por las partículas del sistema, aparece la dificultad[103] derivada del entrelazamiento del tiempo y el espacio señalado en la sección 7.2 sobre la relatividad del espacio físico.
La física relativista vincula las coordenadas de tiempo-espacio de un evento y las componentes de una cantidad física (como la cuadrivelocidad, el cuadrivector de energía-cantidad de movimiento o el cuadrivector potencia-fuerza) asignados a ese evento en un marco inercial K, con las coordenadas de ese mismo evento y las componentes de esas mismas cantidades físicas asignadas a ese mismo evento pero referidas a otro marco inercial K'. Ese vínculo se establece en relatividad restringida utilizando las transformaciones de Lorentz.
Pero si la cantidad física no tiene un carácter local, como por ejemplo no lo tienen las cantidades aditivas tales como el volumen de un cuerpo material, la cantidad de movimiento de un sistema de partículas, o su momento angular, es necesario definir cómo se va a manejar el problema de la simultaneidad de eventos que acontecen en puntos diferentes, cuando se busca relacionar los valores de la cantidad en cuestión referidos a dos marcos inerciales diferentes.

Una posibilidad es trabajar con el mismo conjunto de eventos, aceptando que las cantidades locales que se adicionan y que aparecen como simultáneas para un observador, no sean necesariamente simultáneas para otro observador. Esto permite relacionar las cantidades mediante transformaciones de Lorentz, porque ambos observadores están trabajando con el mismo conjunto de eventos: se puede denominar procedimiento covariante.

Otra posibilidad es decidirse a trabajar con eventos simultáneos en el marco de referencia de cada uno de los observadores, aceptando que a pesar de que los dos se ocupan del mismo sistema (el mismo tubo de universo en el espacio de Minkowski) ya no trabajan con los mismos eventos, y por tanto al no tratarse de los mismos eventos no

---

[102] Dos enfoques distintos y complementarios de la mecánica de fluidos relativista pueden encontrarse en el capítulo 8 del libro de J. L. Synge "*Relativity: the special theory*", North-Holland, Amsterdam, 2ª edición, 1965, y en el capítulo 15 del libro de L. Landau y E. Lifchitz "*Mecánica de fluidos*", Reverté, Barcelona, 1984. El primero es más riguroso y completo en el planteo matemático pero, a diferencia del segundo, no aborda los fenómenos disipativos.

[103] T. Nakamura "Covariant thermodynamics of an object with finite volume", Physics letters A, Vol. 352, 2006, pp. 175-177.

pueden conectar sus componentes mediante transformaciones de Lorentz. Como consecuencia las cantidades físicas no locales en general ya no van a ser las mismas, en el sentido de la física relativista, para los dos observadores: este segundo procedimiento no es covariante y puede conducir a resultados aparentemente paradójicos cuando no se lo distingue del primero[104].
En particular, si siguiendo el segundo procedimiento, se construye un nuevo objeto matemático adicionando los cuadrivectores de energía-cantidad de movimiento correspondientes a diferentes elementos de masa de un cuerpo de volumen finito tomados en forma simultánea respecto de cada observador, en general no se obtiene una entidad covariante (es decir, invariante bajo una transformación de Lorentz)[105]. Esta es una de las dificultades que se plantean a propósito de la formulación de una mecánica de fluidos y una termodinámica relativistas.

Hay una forma directa de evitar las dificultades que acarrea el segundo procedimiento, que ha resultado muy exitosa al ser aplicada en mecánica de fluidos relativista[106].
Consiste en considerar un marco inercial K_0 que en ese instante se halla en reposo respecto de un elemento de masa del fluido (se denomina marco localmente en co-movimiento con el fluido).
Si el fluido es ideal, su estado puede describirse mediante dos variables: la densidad local (densidad propia) de energía interna e_0 y la presión p (la presión es un invariante relativista de modo que $p_0 = p$). La velocidad local del fluido es nula.

A partir de estas dos cantidades se construye una versión **local** (en un punto del medio continuo) del cuadrivector energía-cantidad de movimiento que se puede referir a cualquier marco inercial K.
Se trata de un objeto matemático que respecto a cada marco de referencia se caracteriza por 16 componentes, organizadas en un arreglo matricial simétrico de 4×4.
Es el denominado **tensor de energía-impulsión**. Las componentes $T_{i,j}$ del tensor ($i, j = 0,1,2,3$ donde el sub-índice 0 alude al tiempo asociado con la coordenada $x_0 = c \cdot t$ y los sub-índices 1, 2, 3 se refieren a las componentes $x_1 = x, x_2 = y, x_3 = z$ de un sistema de coordenadas cartesianas ortogonales) se modifican en forma covariante al cambiar el marco inercial (el tensor mantiene su estructura inalterada bajo las transformaciones de Lorentz).
Por ejemplo, para la velocidad local $\vec{v}(t,\vec{r})$ de una partícula de fluido sus componentes cartesianas se representan ahora mediante v_1, v_2 y v_3.

---

[104] A. Gamba "Physical quantities in different reference systems according to relativity", American Journal of Physics, Vol. 35, 1967, pp.83-89.

[105] No obstante, Moller, en la última edición de su libro "*The theory of relativity*", Oxford University Press, Londres, 1972, señala que para un **cuerpo aislado** se puede construir un cuadrivector energía-cantidad de movimiento. Tanto en mecánica de fluidos como en termodinámica es necesario considerar cuerpos en interacción con el medio que los rodea. Por ejemplo cuerpos sometidos a presión, es decir, a un flujo de cantidad de movimiento procedente de su ambiente.

[106] Ver el capítulo 15 de L. Landau y E. Lifchitz "*Mecánica de fluidos*", Reverté, Barcelona, 1984.

Si introducimos el valor local de la cuadrivelocidad relativa $\frac{1}{c}\left(\tilde{U}\right)_K = \left(\gamma, \gamma \cdot \frac{1}{c}\vec{v}\right)_K$ del fluido respecto de un marco K, las componentes del tensor de energía-impulsión se definen así, siendo $\gamma^2 = \frac{1}{1-\frac{v^2}{c^2}}$: $\quad T_{0,0} = \left(e_0 + p\right)\cdot\gamma^2 - p = \gamma^2 \cdot \left(e_0 + \frac{v^2}{c^2}\cdot p\right)$

$$T_{0,j} = T_{j,0} = \left(e_0 + p\right)\cdot\gamma^2 \cdot v_j \qquad j = 1,2,3 \qquad T_{i,j} = \left(e_0 + p\right)\cdot\gamma^2 \cdot v_i \cdot v_j + p \cdot \delta_{i,j} \qquad i, j = 1,2,3$$

Como es usual, $\delta_{i,j}$ es igual a 1 cuando $i = j$ y es igual a cero cuando $i \neq j$.

Constituye una generalización covariante del cuadrivector de energía-cantidad de movimiento, pero restringida a un punto del medio.

A partir de las componentes del tensor de energía-impulsión se pueden escribir en forma covariante dos ecuaciones del movimiento para el fluido perfecto en ausencia de fuerzas de volumen (como las fuerzas electromagnéticas sobre un fluido con carga eléctrica que se consideran en la magneto-hidrodinámica).

La ecuación de conservación de la energía: $\frac{\partial T_{00}}{\partial x_0} + \frac{\partial T_{01}}{\partial x_1} + \frac{\partial T_{02}}{\partial x_2} + \frac{\partial T_{03}}{\partial x_3} = 0$

Las ecuaciones de conservación de la cantidad de movimiento (impulsión):

$$\frac{\partial T_{i0}}{\partial x_0} + \frac{\partial T_{i1}}{\partial x_1} + \frac{\partial T_{i2}}{\partial x_2} + \frac{\partial T_{i3}}{\partial x_3} = 0 \qquad i = 1,2,3$$

Ambas ecuaciones se pueden resumir en una única ley de conservación: $\sum_{k=0}^{3} \frac{\partial T_{i,k}}{\partial x_k} = 0$

El cuadrivector $\sum_{k=0}^{3} \frac{\partial T_{i,k}}{\partial x_k}$ se conoce como **cuadridivergencia** del tensor de energía-impulsión.

Para completar las ecuaciones que describen el comportamiento relativista del fluido es necesario añadir una ecuación de conservación del número de partículas y una relación entre la densidad de energía y la presión.

En forma análoga a la utilizada para definir el tensor de energía-impulsión, se define un cuadrivector local denominado **corriente de partículas**:

$$\left(\tilde{J}\right)_K = n_0 \cdot \left(\frac{1}{c}\cdot\left(\tilde{U}\right)\right)_K = \left(n_0 \cdot \gamma, n_0 \cdot \gamma \cdot \frac{1}{c}\cdot\vec{v}\right)_K$$

La variable $n_0\left(t, \vec{r}\right)$ que aparece en ambas componentes, temporaloide $n_0 \cdot \gamma$ y espacialoide $n_0 \cdot \gamma \cdot \frac{1}{c}\cdot\vec{v}$ de la corriente de partículas, es por definición un escalar (invariante relativista): es el número de partículas por unidad de volumen determinado respecto al marco K_0, es decir, la densidad propia de partículas.

Entonces la ley de conservación del número de partículas se escribe así:

$$\frac{\partial}{\partial x_0}\left(n_0 \cdot \gamma\right) + \frac{\partial}{\partial x_1}\left(n_0 \cdot \gamma \cdot \frac{1}{c}\cdot v_1\right) + \frac{\partial}{\partial x_2}\left(n_0 \cdot \gamma \cdot \frac{1}{c}\cdot v_2\right) + \frac{\partial}{\partial x_3}\left(n_0 \cdot \gamma \cdot \frac{1}{c}\cdot v_3\right) = 0$$

Cuando las velocidades asociadas al movimiento macroscópico del fluido, es decir las magnitudes de los valores del campo $\vec{v}\left(t, \vec{r}\right)$ (cuyas componentes cartesianas son $v_x\left(t, \vec{r}\right)$,

$v_y(t,\vec{r})$ y $v_z(t,\vec{r})$) son pequeñas respecto de c, las ecuaciones del movimiento se reducen a las siguientes:

$$\rho_0 \cdot \left(1+\frac{p}{\rho_0 \cdot c^2}\right) \cdot \frac{d\vec{v}}{dt} = -\nabla p \qquad \frac{\partial \rho_0}{\partial t} + \vec{v} \bullet (\nabla \rho_0) + \rho_0 \cdot \left(1+\frac{p}{\rho_0 \cdot c^2}\right) \cdot \nabla \bullet \vec{v} = 0$$

En estas ecuaciones introdujimos la densidad de masa propia ρ_0 asociada a la energía interna propia del fluido: $e_0 = \rho_0 \cdot c^2$ Como siempre, ∇p y $\nabla \rho_0$ representan los campos vectoriales gradiente, cuyas componentes cartesianas son las derivadas parciales de los campos escalares $p(t,\vec{r})$ y $\rho_0(t,\vec{r})$ respecto de las coordenadas x, y, z.

Por otra parte, $\frac{d\vec{v}}{dt}$ representa, como es usual en mecánica de fluidos clásica, la derivada siguiendo el movimiento $\frac{d\vec{v}}{dt} = \frac{\partial \vec{v}}{\partial t} + (\vec{v} \bullet \nabla)\vec{v}$ mientras que $\nabla \bullet \vec{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}$ es la divergencia del campo de velocidades del fluido y el operador $(\vec{v} \bullet \nabla)$ se expresa en coordenadas cartesianas así: $(\vec{v} \bullet \nabla) = v_x \cdot \frac{\partial}{\partial x} + v_y \cdot \frac{\partial}{\partial y} + v_z \cdot \frac{\partial}{\partial z}$.

Si además de tener velocidades macroscópicas pequeñas respecto de la velocidad de la luz, se tienen presiones p pequeñas respecto de $\rho_0 \cdot c^2$, resultan las ecuaciones clásicas de la dinámica de un fluido ideal en ausencia de fuerzas de volumen:

$$\rho_0 \cdot \frac{d\vec{v}}{dt} = -\nabla p \qquad \frac{\partial \rho_0}{\partial t} + \vec{v} \bullet (\nabla \rho_0) + \rho_0 \cdot \nabla \bullet \vec{v} = 0$$

La relación constitutiva entre la energía interna e_0 y la presión $p = p_0$ conecta la mecánica de fluidos con la termodinámica, debido a que cabe esperar que en esa relación aparezca la temperatura local T del fluido. Nos deja en medio del problema de la termodinámica relativista y la discusión acerca de las leyes de transformación de la temperatura T y el calor Q al cambiar de marco de referencia.

Casi desde que comenzaron los trabajos tendientes a hallar una formulación covariante para la termodinámica hubo un cierto consenso acerca de que la entropía es un invariante relativista[107].

Como consecuencia de la relación $T \cdot dS = \delta Q$ que conecta una variación de la entropía (variable de estado) con la temperatura y con una pequeña cantidad de calor δQ intercambiado entre el sistema y su ambiente, la temperatura absoluta y el calor deben transformarse igual al pasar de un marco de referencia a otro.

Desde el trabajo pionero de K. von Monsegeil (un alumno de doctorado de Planck fallecido prematuramente en un accidente) en 1907 hasta el replanteo debido a H. Ott en 1963, por lo general se admitía que la temperatura T y el calor intercambiado Q referidos a un marco K respecto del cual el sistema termodinámico se encontraba en movimiento con velocidad v se relacionan con los valores propios T_0 y Q_0 de esas

[107] R. Tolman "*Relativity, thermodynamics and cosmology*", Dover, N. Y, 1987, pp. 157-158.

mismas variables referidas a un marco en co-movimiento con el sistema por las fórmulas: $T = \frac{1}{\gamma} \cdot T_0$ y $Q = \frac{1}{\gamma} \cdot Q_0$

Por su parte Ott inició formalmente una segunda corriente de opinión que sostiene que las fórmulas correctas son: $T = \gamma \cdot T_0$ y $Q = \gamma \cdot Q_0$.

En 1969, Van Kampen inició una tercera corriente de opinión proponiendo que tanto la temperatura absoluta como el calor intercambiado son invariantes relativistas: $T = T_0$ y $Q = Q_0$

Otros han propuesto que la temperatura es invariante mientras que el calor no lo es, o han propuesto fórmulas para expresar la covariancia diferentes a las expuestas previamente, o han llegado a la conclusión que el intento de formulación covariante de la termodinámica carece de sentido[108].

Hay dos formas de evitar el problema. Una es suponer para un fluido ideal una relación entre e_0 y p solamente. La otra consiste en trabajar siempre con variables termodinámicas propias, medidas en un marco en co-movimiento local con el fluido perfecto. Así, en la relación constitutiva aparece la temperatura propia T_0. Como en nuestro caso el interés por la termodinámica relativista se vincula con la construcción de modelos cosmológicos que se discute en el capítulo 12, adoptaremos este último punto de vista.

8 .8 Las relaciones epistemológicas entre las teorías de la física clásica luego de ser aceptada la teoría restringida de la relatividad

Para finalizar este capítulo, consideremos desde un punto de vista epistemológico, las nuevas relaciones entre las teorías físicas cuando aparece la física relativista en 1905, y hasta 1915, antes de aparecer la teoría generalizada.

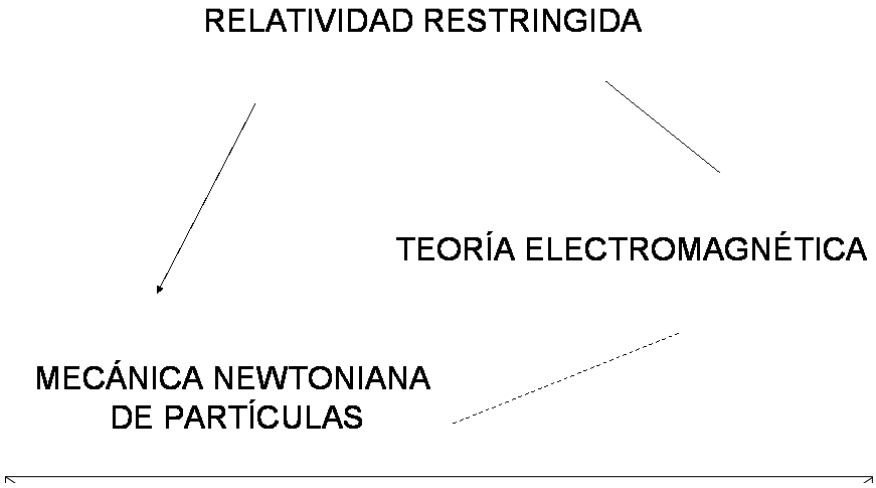
El siguiente esquema describe las relaciones de dominio, compatibilidad o inconsistencia que se presentan entre la teoría restringida de la relatividad, la mecánica newtoniana de partículas y la teoría electromagnética.

La flecha indica el dominio de la teoría restringida sobre la mecánica newtoniana. Ésta se obtiene como caso límite de la dinámica relativista haciendo tender a infinito la velocidad de la luz.

El segmento de línea continua que conecta la relatividad restringida con la teoría electromagnética representa la compatibilidad entre ambas teorías, sin que una domine sobre la otra.

La línea de trazo punteado representa la inconsistencia entre la mecánica newtoniana de partículas y la teoría electromagnética.

[108] Ver H. Callen y G. Horwitz "Relativistic thermodynamics", American Journal of Physics, Vol. 39, 1971, pp. 938-947 y M. Requardt "Thermodynamics meets special relativity- or what is real in physics?" arXiv:0801.2639v1 (gr-qc), 2008.



No se ha incluido la termo-mecánica de medios continuos, puesto que ni sus versiones newtonianas ni sus versiones relativistas han alcanzado el grado de completitud (y aceptación, en el caso relativista) alcanzado por la mecánica de partículas, tanto newtoniana como relativista.
La teoría de la gravitación de Newton permaneció como una teoría de acción a distancia, inexplicada e incompatible con la física relativista, hasta que en 1915 fue superada por una nueva teoría, relativista, del campo gravitatorio.

El concepto de límite clásico, aceptado por los físicos, según el cual las fórmulas de la mecánica de Newton se obtienen como aproximación a partir de las fórmulas de la mecánica relativista cuando las velocidades de los cuerpos son pequeñas respecto de la velocidad de la luz, ha sido objetado por algunos filósofos de la ciencia, desde Thomas Kuhn hasta Paul Feyeraband.
Estos pensadores enfatizan la inconmensurabilidad entre una y otra mecánica, y en general entre teorías separadas por un cambio de paradigma[109]: "Con el paso de una teoría a la que le sigue, las palabras transforman sus significados o las condiciones de su aplicabilidad de una forma muy sutil. Aunque en buena medida se utilizan los mismos signos antes y después de la revolución- por ejemplo, fuerza, masa, elemento, compuesto, célula- se ha modificado de alguna manera el tipo y el modo de su

[109] En líneas generales se pueden distinguir dos etapas que se repiten durante el desarrollo histórico de las teorías físicas. En una de esas etapas la teoría dominante ha alcanzado un grado de completitud tal que no puede ser significativamente mejorada introduciendo pequeñas modificaciones. Se ha alcanzado el estadio de una teoría cerrada en el sentido definido por Heisenberg en su artículo *"Der Begriff abgeschlossene Theorie in der moderne Naturwissenschaft",* Dialectica, vol. 2, 1948, pp. 331-336. Corresponde a lo que Kuhn (*"La estructura de las revoluciones científicas",* FCE, Madrid y México, 2ª edición, 1976) propuso en 1962 denominar etapa de ciencia normal, gobernada por un paradigma característico, que en ese momento resulta muy fructífero. En la siguiente etapa aparecen dificultades que no pueden resolverse en el marco de la teoría establecida. El paradigma es cuestionado y finalmente resulta sustituido por otro en el marco de la elaboración de una nueva teoría que termina por sustituir a la anterior. Este proceso de cambio de paradigma es denominado, a propuesta de Kuhn, revolución científica. Los límites de aplicabilidad de la vieja teoría solo se comprenden en el marco de la teoría nueva. Esta se cierra a su vez dando paso a una nueva etapa de ciencia normal en torno al correspondiente paradigma.

aplicación a la naturaleza. Por ello decimos que las teorías sucesivas son inconmensurables."[110]
Cuando varía lo suficiente el significado de los términos, la traducción de uno de los lenguajes teóricos al otro ya no resulta posible y como consecuencia de ello los contenidos de las teorías no pueden ser comparados.
Por ejemplo, la palabra masa en la mecánica newtoniana no tiene el mismo significado que en mecánica relativista, puesto que en la primera se encuentra sujeta a una ley de conservación mientras que en la segunda, su homóloga, la masa inercial, puede desaparecer siempre que se conserve la energía[111].
Lo mismo ocurre con el tiempo y el espacio, independientes entre sí en mecánica newtoniana e interrelacionados en mecánica relativista.
Un problema similar se plantea a propósito del concepto de fuerza. Introducido originalmente en el contexto de la estática, es posible trasladarlo a la mecánica relativista mientras se añada la restricción de equilibrio de fuerzas.
No obstante, extendido a la dinámica, el concepto newtoniano de fuerza implica la tercera ley de Newton: igualdad de acción y reacción entre cuerpos situados en puntos del espacio distantes, en cada instante de tiempo. Para que la tercera ley se verifique respecto de todo marco de referencia posible, como se asume en mecánica newtoniana, es necesario que se verifique la simultaneidad absoluta de eventos en puntos diferentes. En consecuencia el concepto newtoniano de fuerza no resulta traducible al lenguaje de la teoría de la relatividad. Es incompatible con el concepto de fuerza relativista, excepto en un único caso: si nos restringimos a las fuerzas de contacto, como la presión en un medio continuo, puesto que entonces acción y reacción se refieren siempre al mismo punto del espacio.

Pero la traducibilidad estricta no es un requisito necesario para poder comparar teorías en física. Como señala Rivadulla, en física lo decisivo para poder comparar dos teorías, por lo demás inconmensurables, es la identificación de homologías entre los términos propios de los lenguajes de esas teorías y el balance predictivo de cada una, es decir, la capacidad de dar cuenta de las relaciones empíricas conocidas y de predecir nuevas relaciones.[112]
La masa en mecánica newtoniana es homóloga a la masa en reposo de la dinámica relativista, los componentes espacialoides del cuadrivector de energía-cantidad de movimiento son homólogos a la cantidad de movimiento de la mecánica clásica, y las componentes espacialoides en la igualdad entre el cuadrivector de potencia-fuerza y la derivada respecto del tiempo propio del cuadrivector de energía-cantidad de movimiento, son homólogas a las componentes escalares de la igualdad vectorial que expresa la segunda ley de Newton como relación entre la fuerza y la velocidad de

---

[110] T. Kuhn *"La estructura de las revoluciones científicas"*, FCE, Madrid y México, 2ª edición, 1976, sección 6.

[111] Esto último puede acontecer aunque la velocidad de las partículas sea pequeña respecto de la velocidad de la luz, es decir, en el límite newtoniano, como se constata en las reacciones nucleares de fisión del tipo de las que se producen en reactores de investigación o de potencia (R.Suárez Ántola "*La Energía Nuclear: aspectos científicos, técnicos y sociales de la conversión núcleo-eléctrica*", resa.libros@gmail.com, Montevideo, Uruguay, 2009).

[112] A. Rivadulla "*Éxito, razón y cambio en física: un enfoque instrumental en teoría de la ciencia*", Trotta, Madrid, 2004, capítulo 4.

cambio en la cantidad de movimiento.
Estas homologías permiten aplicar el principio de consistencia planteado en 8.1, según el cual la mecánica basada en la teoría restringida de la relatividad, que pretende reemplazar a la mecánica de Newton, debe dar cuenta de todas las predicciones exitosas de esta última.

9 Espacios curvos

"Gauss (1827) hizo el primer estudio sistemático de las formas diferenciales cuadráticas en sus "Disquisitiones generales circa superficies curvas", en las que el tema principal es la curvatura de las superficies… Dando el paso final en esta dirección, Riemann, en una de las contribuciones más prolíficas hechas a la geometría, pasó de inmediato a la forma diferencial cuadrática general en n variables, con coeficientes variables, en su "Über die Hypothesen welche der Geometrie zu Grunde liegen", 1854. … La matemática de la geometría de Riemann entrelaza dos temas fundamentales: una generalización del teorema de Pitágoras para cualquier espacio (variedad numérica) de n dimensiones y la curvatura en esos espacios. …Hizo entonces la notable conjetura de que su nueva métrica reduciría a geometría pura las cuestiones relativas al universo material y a las fuerzas de enlace que lo mantendrían unido."
E. Bell "The development of mathematics", Mc Graw-Hill, New York, 1940.

Debido a la igualdad entre la masa inercial y la masa gravitatoria, los movimientos en un campo gravitatorio poseen la misma aceleración en el mismo punto y en el mismo instante. En consecuencia, siempre que las condiciones iniciales hayan sido las mismas, todos los cuerpos de prueba se mueven de la misma forma, siguiendo la misma trayectoria, independientemente de sus masas o cargas eléctricas. Esto no se observa en el caso de los movimientos en un campo electromagnético, debido a que la aceleración depende ahora del cociente entre la carga eléctrica y la masa inercial, cociente que en general es diferente para cuerpos diferentes.
La idea subyacente a la teoría generalizada de la relatividad es que la igualdad entre la masa inercial y la masa gravitatoria se hace comprensible si el movimiento en un campo gravitatorio se concibe como un movimiento libre de fuerzas en un tiempo-espacio curvo.

9.1 Teoría de superficies y geometrías intrínsecas[113]

La introducción de la idea de un espacio curvo conduce a la consideración de las geometrías no euclidianas, y en particular a las geometrías de Riemann en el caso de la teoría de la relatividad. Antes de considerar estas geometrías, es conveniente revisar la teoría clásica de superficies en el espacio euclidiano E_3 y el concepto de geometría intrínseca a una superficie en dicho espacio.

9.1.1 Curvas, superficies y curvaturas en E_3.

Comencemos por revisar el concepto de curvatura de una curva plana, dada por una función regular $y = f(x)$.

[113] Una presentación resumida de los temas fundamentales de geometría diferencial clásica se puede encontrar, por ejemplo, en el libro de H. Coxeter "*Introduction to geometry*", Parte 4, Wiley, New York, 1980, y en el capítulo 14 del volumen 2 del libro de H. Ewes "Estudio de las geometrías", UTHEA, México, 1969. Un tratamiento más detallado y avanzado se encuentra en los libros de D. Struik "Geometría diferencial clásica", Aguilar, Madrid, 1966, y de H. Guggenheimer "Differential geometry", Dover, N.Y., 1977.

En términos que apelan a la intuición, podemos decir que la curvatura se introduce como una medida de la rapidez con que rota la recta tangente a medida que pasamos de un punto de la curva a otro próximo.
Para que la curvatura sea independiente de la orientación de la curva en el plano, se define como el límite del cociente entre el ángulo $\Delta\alpha$ que forman las rectas tangentes en dos puntos y la longitud Δs del arco de curva que tiene a esos puntos como extremos, manteniendo un punto fijo y aproximando el otro: $k = \lim_{\Delta s \to 0}\left(\frac{\Delta\alpha}{\Delta s}\right)$

Introduciendo la abscisa x para expresar tanto la longitud de arco $s(x)$ como el ángulo $\alpha(x)$ entre la recta tangente a la curva en el punto $(x, f(x))$ y el eje x, y representando las derivadas primera y segunda por f' y f'' se obtiene la fórmula bien conocida[114]:

$$k(x_0) = \frac{f''(x_0)}{\left(1+[f'(x_0)]^2\right)^{\frac{3}{2}}}$$

Se puede definir un radio de curvatura mediante la relación $R(x_0) = \frac{1}{|k(x_0)|}$.

Debido a cómo se definen, resulta evidente que tanto la curvatura como el radio de curvatura son propiedades locales.
Desarrollemos en serie de Taylor la función $y = f(x)$ en torno a un punto de abscisa x_0 hasta el orden dos inclusive, representando con $o(\ )$, como es usual, un infinitésimo de orden superior al de su argumento:

$$y = f(x) = f(x_0) + f'(x_0)\cdot(x - x_0) + \frac{1}{2}\cdot f''(x_0)\cdot(x - x_0)^2 + o\left((x - x_0)^2\right)$$

Rotemos los ejes de coordenadas y traslademos el origen al punto $(x_0, f(x_0))$, de modo que el nuevo eje de abscisas $\tilde{x}$ coincida con la recta tangente a la curva en el punto $(x_0, f(x_0))$ y el nuevo eje de ordenadas $\tilde{y}$ mida el apartamiento de la curva de la mencionada recta tangente. En los nuevos ejes la misma curva se representa mediante la función $\tilde{y} = g(\tilde{x})$ cuyo desarrollo de Taylor es ahora:

$$\tilde{y} = g(\tilde{x}) = \frac{1}{2}\cdot g''(0)\cdot \tilde{x}^2 + o\left(\tilde{x}^2\right)$$

La curvatura de la curva plana se calcula empleando las nuevas coordenadas aplicando la fórmula de la curvatura para este caso: $k(0) = \frac{g''(0)}{\left(1+[g'(0)]^2\right)^{\frac{3}{2}}} = g''(0) = \lim_{\tilde{x}\to 0}\left(\frac{2\cdot\tilde{y}}{\tilde{x}^2}\right)$

Como la nueva ordenada mide el apartamiento de la curva respecto de la recta tangente en el punto considerado (el punto cuyas coordenadas originales son $(x_0, f(x_0))$), la

[114] Teniendo en cuenta que $\alpha(x) = arctg(f'(x))$ y que $\frac{d}{dx}s(x) = \sqrt{1+[f'(x)]^2}$

fórmula aplicada a la representación de la curva en las nuevas coordenadas muestra que la curvatura describe la rapidez con que la curva abandona la tangente[115] (Figura 9.1).

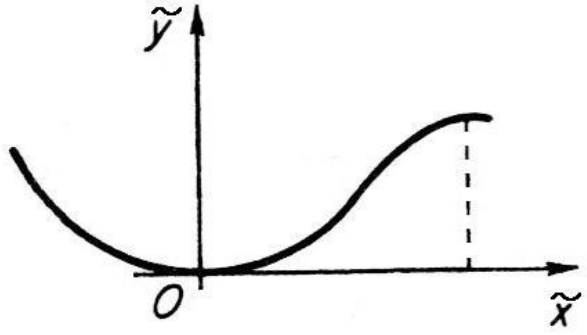


Figura 9.1 Apartamiento de la curva respecto de la tangente

En el caso de una superficie orientable (bilátera) regular Σ, parecería razonable describir su curvatura en un punto dado mediante la rapidez con que dicha superficie se separa de su plano tangente en el punto.
Como en direcciones diferentes la superficie puede separase de su plano tangente con diferente rapidez[116], una forma de implementar esta idea es asignarle las curvaturas de curvas planas contenidas en la superficie y que pasan por el punto considerado en todas las direcciones posibles. Para ello, se construye la recta normal al plano tangente en el punto, la denominada normal a la superficie en el punto. Luego se construyen todos los planos que pasan por esa normal (planos normales) y se intersecan con Σ. A cada plano se le puede asignar un ángulo ϕ, formado por dos rectas dirigidas, situadas en el plano tangente: una de ellas es fija, y la otra es la intersección del plano normal con el plano tangente. Se obtiene el conjunto de curvas planas (conocidas como secciones normales) que se buscaba: pasan por el punto en todas las direcciones posibles y a cada una se le puede calcular la curvatura $k_N(\phi)$ en el punto. Pero la dependencia entre la curvatura y la dirección de la curva no es arbitraria: la curvatura de una superficie en un punto se puede caracterizar por dos parámetros.
Para verlo, consideremos una superficie Σ dada por la función regular $z = f(x, y)$, donde x, y, z son coordenadas cartesianas ortogonales.

---

[115] En un entorno del origen la curva $\tilde{y} = g(\tilde{x})$ puede aproximarse por la parábola $\frac{1}{2} \cdot g''(0) \cdot \tilde{x}^2$ con un error que tiende a 0 como $\tilde{x}^3$ cuando $\tilde{x}$ tiende a cero. Esta parábola, así como la curva $\tilde{y} = g(\tilde{x})$ misma, se pueden aproximar a su vez mediante la denominada circunferencia osculadora $(\tilde{y} - \tilde{y}_c) + \tilde{x}^2 = R^2$ si se eligen adecuadamente $\tilde{y}_c$ y R. Desarrollando la fórmula de la circunferencia y reordenando se obtiene: $\tilde{y} = \frac{1}{2 \cdot \tilde{y}_c} \cdot \tilde{x}^2 + \frac{(\tilde{y}_c^2 - R^2) + \tilde{y}^2}{2 \cdot \tilde{y}_c}$ Para que la circunferencia aproxime a la parábola debe verificarse: (1) $\tilde{y}_c^2 = R^2$ (o sea $\tilde{y}_c = \pm R$), y (2) $\frac{1}{\tilde{y}_c} = g''(0) = k(0)$ Entonces $R = \pm \frac{1}{k(0)}$ es el **radio de curvatura** de $\tilde{y} = g(\tilde{x})$ en el origen, y por tanto el radio de curvatura de la curva $y = f(x)$ en el punto $(x_0, f(x_0))$.

[116] Para verlo basta pensar en lo que ocurre en las proximidades de un punto ubicado sobre una superficie cilíndrica cuando se construye el plano tangente por ese punto.

En un punto (x_0, y_0, z_0) con $z_0 = f(x_0, y_0)$ el plano tangente viene dado por la función lineal $p(x,y) = f(x_0,y_0) + \frac{\partial f(x_0,y_0)}{\partial x}\cdot(x-x_0) + \frac{\partial f(x_0,y_0)}{\partial x}\cdot(y-y_0)$

El apartamiento de la superficie respecto del plano tangente puede expresarse mediante un reordenamiento del desarrollo de Taylor de $z = f(x,y)$:

$$f(x,y) - p(x,y) =$$
$$\frac{1}{2}\left[\frac{\partial^2 f(x_0,y_0)}{\partial x^2}\cdot(x-x_0)^2 + 2\frac{\partial^2 f(x_0,y_0)}{\partial x\cdot\partial y}\cdot(x-x_0)\cdot(y-y_0) + \frac{\partial^2 f(x_0,y_0)}{\partial y^2}\cdot(y-y_0)^2\right] + o\left((x-x_0)^2 + (y-y_0)^2\right)$$

Trasladamos el origen de coordenadas al punto (x_0, y_0, z_0) y rotemos los ejes de modo que los nuevos ejes $\tilde{x}, \tilde{y}$ se encuentren en el plano tangente y el nuevo eje $\tilde{z}$ coincida con la normal a la superficie en (x_0, y_0, z_0).

Es posible ajustar la posición de los ejes $\tilde{x}, \tilde{y}$ en el plano tangente para que se anule en el origen la derivada cruzada $\frac{\partial g(\tilde{x},\tilde{y})}{\partial x\cdot\partial y}$ de la función $\tilde{z} = g(\tilde{x},\tilde{y})$ que representa a Σ en las nuevas coordenadas. Entonces una medida del apartamiento de la superficie respecto de su plano tangente viene dada por:

$$\tilde{z} = g(\tilde{x},\tilde{y}) = \frac{1}{2}\left[\frac{\partial^2 g(0,0)}{\partial \tilde{x}^2}\cdot\tilde{x}^2 + \frac{\partial^2 g(0,0)}{\partial \tilde{y}^2}\cdot\tilde{y}^2\right] + o\left(\tilde{x}^2 + \tilde{y}^2\right)$$

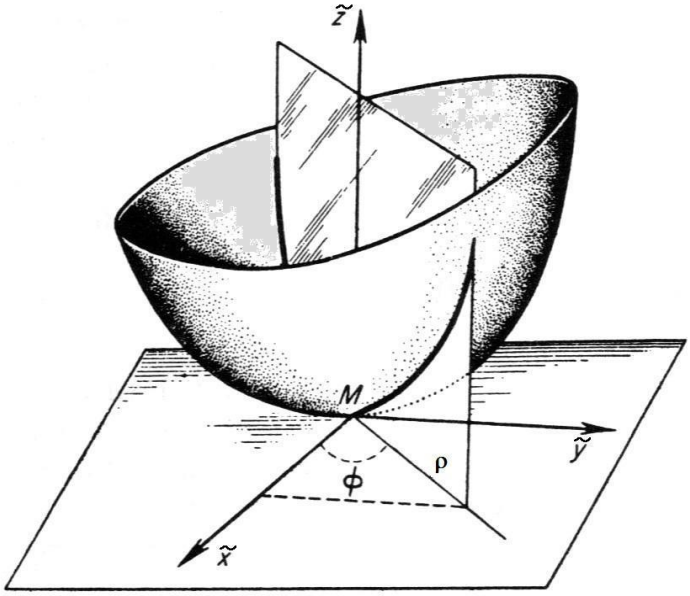


Figura 9.2 Apartamiento de la superficie respecto del plano tangente.

Si el ángulo ϕ de un plano normal se mide respecto del eje $\tilde{x}$, y se introduce un sistema de coordenadas polares ρ, ϕ, como en ese caso $\tilde{x} = \rho\cdot\cos\phi$ y $\tilde{y} = \rho\cdot sen\phi$, la curva plana determinada por la intersección del plano normal con la superficie viene dada por:

$$\tilde{z} = h_N(\rho) = g(\rho\cdot\cos\phi, \rho\cdot sen\phi) = \frac{\rho^2}{2}\left[\frac{\partial^2 g(0,0)}{\partial \tilde{x}^2}\cdot\cos^2\phi + \frac{\partial^2 g(0,0)}{\partial \tilde{y}^2}\cdot sen^2\phi\right] + o\left(\rho^2\right)$$

Entonces teniendo en cuenta que la curvatura en el origen viene dada por $k_N(\phi) = \lim_{\rho\to 0}\left(\frac{2\cdot\tilde{z}}{\rho^2}\right)$ para $\tilde{z} = h_N(\rho)$, resulta el teorema de curvatura debido a Euler:

$$k_N(\phi) = k_1 \cdot \cos^2\phi + k_2 \cdot sen^2\phi$$

En esta fórmula, por definición, $k_1 = \dfrac{\partial^2 g(0,0)}{\partial \tilde{x}^2}$ y $k_2 = \dfrac{\partial^2 g(0,0)}{\partial \tilde{y}^2}$.

Entonces en un entorno del punto (x_0, y_0, z_0) (o sea del punto $\tilde{x} = \tilde{y} = \tilde{z} = 0$) la superficie se puede representar así: $\tilde{z} = k_1 \cdot \tilde{x}^2 + k_2 \cdot \tilde{y}^2 + o(\tilde{x}^2 + \tilde{y}^2)$

Dependiendo del comportamiento de la superficie en un entorno del punto (x_0, y_0, z_0) k_1 y k_2 pueden tener el mismo signo o signo opuesto, o bien anularse uno o ambos números.

Si tienen el mismo signo la superficie se dispone localmente de un mismo lado del plano tangente, como un cuenco: el punto se denomina **elíptico**.

Si poseen signo opuesto la superficie se ubica localmente a ambos lados del plano tangente, al cual atraviesa, como una silla de montar: el punto se denomina **hiperbólico**.

Si uno de estos números es cero y el otro no lo es, el punto se denomina **parabólico**. Si ambos son iguales se trata de un punto **umbílico**.

Si son nulos se denomina punto plano y la determinación del comportamiento local de la superficie requiere un estudio basado en los términos superiores al segundo en el desarrollo de Taylor de la función $\tilde{z} = g(\tilde{x}, \tilde{y})$.

Si k_1 y k_2 toman valores diferentes, digamos si $k_1 > k_2$, el examen de la fórmula $k_N(\phi) = k_1 \cdot \cos^2\phi + k_2 \cdot sen^2\phi$ muestra que $k_1 \geq k_N(\phi) \geq k_2$ para toda curva contenida en la superficie y en un plano normal.

Las curvaturas extremas se alcanzan en dos direcciones ortogonales, para $\phi = 0$ y para $\phi = \dfrac{\pi}{2}$, que se denominan **direcciones principales**. Las curvas planas cuyas curvaturas son k_1 y k_2 se denominan **secciones principales**, y esas curvaturas se denominan **curvaturas principales.** Son los dos parámetros que en general determinan la curvatura $k_N(\phi)$ de cualquier sección normal. A menos de un signo, esas curvaturas son recíprocas de los denominados radios de curvatura principales $\dfrac{1}{R_1}$ y $\dfrac{1}{R_2}$ de esa superficie en el punto considerado.

Si cortamos la superficie por un plano que pasa por el punto P considerado pero no pasa por la normal a la superficie en ese punto, tenemos una sección plana, pero que ya no es una sección normal.

Si θ es el ángulo que el plano oblicuo forma con la normal, la distancia entre la curva plana oblicua y el plano tangente, **determinada sobre el plano oblicuo de la curva**, es $h_\theta(\rho) = \dfrac{|h_N(\rho)|}{\cos\theta}$. Cuanto más oblicuo es el plano, mayor es $|h_\theta(\rho)|$ respecto de $|h_N(\rho)|$

Entonces aplicando la fórmula para calcular la curvatura en el punto P, es decir $k(\phi) = \lim\limits_{\rho \to 0}\left(\dfrac{2 \cdot h(\rho)}{\rho^2}\right)$, obtenemos el teorema debido a Meusnier: $k(\phi) = \dfrac{|k_N(\phi)|}{\cos\theta}$

Se ve que la magnitud de la curvatura de la sección normal en un punto es menor que las curvaturas de las secciones oblicuas en ese mismo punto y en la misma dirección.

Si la curva Γ es regular pero no está contenida en un plano, para describirla es necesario asignar a cada punto dos curvaturas. Para hacer esto se introduce el concepto de plano

osculador, que contiene a la recta tangente y a una recta normal denominada normal principal, que se ubica en la dirección en la que varía la recta tangente en un entorno del punto. Se define la **curvatura de flexión** κ como una medida siempre positiva, respecto de la longitud de arco correspondiente, de la rapidez con que la proyección de la curva Γ sobre el plano osculador se aparta de la tangente. La dirección del plano osculador en el espacio viene determinada por su correspondiente recta normal. Esta recta, cuando se refiere al punto de la curva considerado, se denomina recta binormal, reservándose el uso de recta normal principal para la recta N contenida en el plano osculador (Figura 9.3).

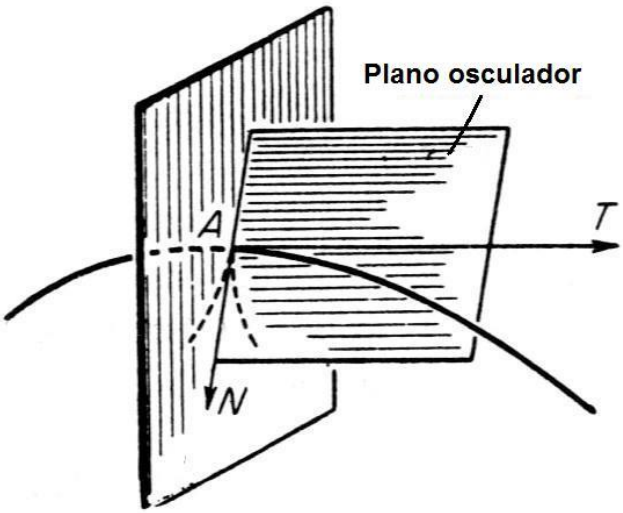


Figura 9.3 Rectas tangente T y normal principal N de una curva no contenida en un plano.

La otra curvatura, denominada **curvatura de torsión** τ, se define como una medida, también respecto de la longitud de arco, de la rapidez con que varía la dirección de los planos osculadores en puntos de la curva próximos al considerado. La curvatura de flexión es la curvatura de la curva plana obtenida proyectando los puntos de la curva Γ sobre el plano osculador. La curvatura de torsión, como su nombre lo indica, mide la torsión, es decir el apartamiento local de Γ respecto de una curva plana. La curvatura de torsión de una curva plana es cero en todos sus puntos.

Consideremos una curva cualquiera contenida en la superficie y que pase por un punto P en el que se quiere estudiar la curvatura. El plano osculador a la curva en el punto P por lo general es oblicuo. La intersección de ese plano con la superficie determina una sección plana cuya curvatura es en valor absoluto igual a la curvatura de flexión de la curva en P. Así pues, si conocemos las curvaturas principales k_1 y k_2 en un punto de una superficie, la curvatura de flexión de cualquier curva en la superficie en ese punto queda determinada por la dirección de su tangente y por el ángulo entre su plano osculador y la normal a la superficie en el punto. Como la curvatura de torsión es independiente de la curvatura de flexión, debe determinarse en forma independiente.

Gauss mostró que es posible definir la curvatura de una superficie en un punto desde un punto de vista diferente al que reduce la cuestión a hallar las curvaturas de las secciones normales por ese punto.

Partió de la definición de curvatura para una curva plana. Como a cada recta tangente le corresponde una recta normal, puede decirse también que la curvatura de una curva plana mide la rapidez de variación de dirección de las rectas normales al pasar de un punto a otro próximo, o dicho de otra forma, mide la divergencia local de las normales a la curva plana.
Supongamos que en el plano de la curva se construye una circunferencia con radio unidad. Entonces a cada punto de la curva le corresponde un punto en la circunferencia, precisamente el extremo del radio que es paralelo a la recta normal al punto de la curva considerado. Al ángulo entre dos normales le corresponde un arco de circunferencia cuya longitud es igual a ese ángulo. La curvatura es el límite del cociente entre la longitud del arco de circunferencia y la longitud del arco de curva correspondiente, cuando esta última longitud tiende a cero de forma tal que un mismo punto de la curva pertenece a todos esos arcos (precisamente el punto en el cual se calcula la curvatura).

Ahora supongamos que tenemos una superficie Σ regular y bilátera en E_3. Consideremos un punto fijo P perteneciente a esa superficie, y una porción $\Delta\Sigma$ de superficie que incluye a P en su interior. Las dimensiones de $\Delta\Sigma$ se tomarán tan pequeñas como pudiera ser conveniente, siempre con el punto P en su interior.
Construimos una esfera E de radio unidad y centro O. Trazamos por el punto O rectas paralelas a las normales a la superficie en cada uno de los puntos de una porción de superficie $\Delta\Sigma$. Entonces a cada punto de $\Delta\Sigma$ le corresponde un punto en la superficie esférica perteneciente a un casquete $\Delta\mathrm{E}$. Obtenemos así una representación esférica de un elemento de superficie. Debemos distinguir entre los dos lados de la superficie, elegir uno de ellos y orientar las rectas normales en consecuencia. Como las normales a la superficie en puntos distintos podrían ser paralelas, algunas porciones de esfera se podrían superponer proviniendo de porciones disjuntas de la superficie. Pero si seleccionamos porciones de superficie lo bastante pequeñas, esta superposición no ocurrirá.
Se selecciona $\Delta\Sigma$ de modo que su frontera es un arco de curva cerrada simple C en Σ. La frontera de $\Delta\mathrm{E}$ es un arco de curva cerrada simple C_0 en E que constituye una imagen esférica de C.
Si orientamos el contorno C en uno de los dos sentidos posibles, el sentido del contorno C_0 queda determinado.
Si su sentido es el mismo que el sentido de C, al área ΔA de $\Delta\Sigma$ y al área $\Delta\Omega$ de su imagen esférica $\Delta\mathrm{E}$ se les asigna el mismo signo. Esto ocurre cuando el punto P es elíptico.
Si el sentido de C_0 resulta el opuesto al de C, al área $\Delta\Omega$ se le asigna diferente signo que el asignado al área ΔA. Esto ocurre cuando el punto P es hiperbólico.
Si se toma el límite del cociente $\dfrac{\Delta\Omega}{\Delta A}$, cuando ΔA tiende a cero de tal modo que el punto P en el que se quiere definir la curvatura pertenezca a todos los elementos de superficie utilizados para calcular el límite, se obtiene la curvatura de Gauss en ese punto: $K(P) = \lim\limits_{\Delta A \to 0}\left(\dfrac{\Delta\Omega}{\Delta A}\right)$
La curvatura K mide la divergencia local de las normales a la superficie.

Esta medida de la curvatura verifica: $K = k_1 \cdot k_2 = \pm \frac{1}{R_1} \cdot \frac{1}{R_2}$ [117]

Además, es una propiedad intrínseca de la superficie[118]: no varía al deformarla sin dilatación ni corte.

Por este motivo a la curvatura de Gauss también se la denomina curvatura intrínseca.

Es una medida de la diferencia entre la geometría intrínseca de la superficie y la geometría plana.

9.1.2 Resumen de algunos resultados sobre geometrías intrínsecas de superficies en E_3.

La consideración del quinto postulado de Euclides (por un punto exterior a una recta dada pasa una y solo una recta paralela a la primera) y la distinción entre geometría como construcción matemática y geometría como modelo del espacio físico, condujo al desarrollo de las geometrías no euclidianas.

Al comienzo de este proceso, Legendre obtuvo un resultado muy importante: demostró que el quinto postulado equivale al teorema que afirma que la suma de los ángulos de un triángulo es igual a π.

Hacia 1816, Gauss ya se había convencido acerca de la imposibilidad de demostrar el quinto postulado, y se encontraba desarrollando la teoría de las superficies embebidas en el espacio euclidiano tridimensional, con ayuda del cálculo infinitesimal.

Caracterizaba las coordenadas cartesianas de los puntos de una superficie mediante funciones del tipo $x = x(u,v)$ $y = y(u,v)$ $z = z(u,v)$, donde el par ordenado (u,v) formado por las variables independientes se pueden interpretar como un punto de una región adecuada en un plano euclidiano. Se denominan parámetros de la representación de la superficie.

Una curva en el espacio se puede caracterizar por funciones del tipo $x = x(u)$ $y = y(u)$ $z = z(u)$. En particular, si en las funciones que determinan los puntos de una superficie se fija un de las variables independientes, haciendo por ejemplo $v = v_0$, se obtiene una curva cuyos puntos se encuentran contenidos en la superficie $x = x(u, v_0)$ $y = y(u, v_0)$ $z = z(u, v_0)$ Una curvas correspondiente a un valor constante de uno de los parámetros se conoce como **línea de coordenada** o **curva coordenada**.

En forma análoga se puede fijar la otra variable independiente, o definir una relación entre ambas variable, determinando así otras tantas curvas incluidas en la superficie.

Se puede construir entonces una red de curvas que cubren toda la superficie y se intersecan entre sí (no siendo en general recto el ángulo definido por la intersección: no son necesariamente familias de curvas ortogonales).

El elemento de línea $ds^2 = dx^2 + dy^2 + dz^2$ que mide la separación de dos puntos en el

---

[117] Si la superficie es esférica y su radio es R_0, resulta que $K = \frac{1}{{R_0}^2}$ es positiva y constante.

[118] La otra curvatura local empleada en la teoría de superficies, la **curvatura media** o curvatura de Germain, es una propiedad de la geometría extrínseca de la superficie, es decir depende de cómo la superficie se encuentra inmersa en el espacio tridimensional. Ver 9.4.1.

espacio euclidiano se puede aplicar a dos puntos de la superficie:

$$ds^2 = dx(u,v)^2 + dy(u,v)^2 + dz(u,v)^2$$

Para dos puntos vecinos Gauss dedujo la expresión del elemento de longitud en función de las coordenadas (u,v) (en 9.4.1 se puede hallar una deducción):

$$ds^2 = E(u,v)\cdot du^2 + 2\cdot F(u,v)\cdot du\cdot dv + G(u,v)\cdot dv^2$$

Aquí u y v son los parámetros de la red de curvas o líneas de coordenadas contenidas en la superficie y tales que por cada punto pasan en general dos de estas curvas.

Ahora bien, la geometría en el plano euclidiano se reduce a estudiar propiedades de líneas y figuras sobre la base de las distancias entre sus puntos.

Gauss de dio cuenta que si se conocían E, F y G como funciones de u y v era posible construir una geometría en la superficie. Para hacerlo se basó en la posibilidad de caracterizar algunos elementos geométricos básicos, como las líneas rectas y los ángulos de la geometría de Euclides en el plano, de tal manera que sea factible identificar los elementos correspondientes en una superficie cualquiera.

Las líneas rectas, y solamente ellas, son las curvas de longitud mínima que pasan por dos puntos dados del plano o del espacio euclidiano.

En una superficie pueden definirse entonces curvas equivalentes a las rectas del espacio euclidiano como las **líneas de menor longitud** que unen dos puntos dados: las **líneas geodésicas.**

Dados tres puntos en una superficie, no pertenecientes a la misma línea geodésica, se construyen las líneas geodésicas que los unen y queda definido el equivalente de un triángulo. Por cada vértice del triángulo pasan dos geodésicas, y las tangentes a esas geodésicas en dicho vértice determinan el correspondiente ángulo del triángulo.

En el caso de una superficie esférica, si α,β,γ son los ángulos de un triángulo contenido en la superficie, si σ es el área de ese triángulo, y si K es la medida (intrínseca) de la curvatura introducida por Gauss se tiene[119]: $\alpha+\beta+\gamma-\pi=\sigma(K+\varepsilon)$ donde ε tiende a cero cuando el triángulo se contrae hacia un punto de la superficie y su área tiende a cero. Entonces, cuanto menor es ese triángulo, más próxima a π se encuentra la suma de sus tres ángulos. En ese sentido, la geometría en una superficie esférica es, localmente, una geometría euclidiana plana.

Reordenando la fórmula para la suma de los ángulos de un triángulo se obtiene una fórmula para la curvatura de Gauss: $K=\dfrac{\alpha+\beta+\gamma-\pi}{\sigma}-\varepsilon$ se ve que la curvatura local (como $K=\dfrac{1}{R_0{}^2}$ el radio de la esfera $R_0=\dfrac{1}{\sqrt{K}}$) se puede estimar sumando los ángulos de un triángulo, restando π y dividiendo el resultado por el área del triángulo, siendo la estimación tanto más aproximada cuanto más pequeño se tome el triángulo, puesto que ε tiende a cero con esa área.

Pero Gauss fue aún más lejos. Comprendió que la geometría bidimensional que se podía construir a partir del elemento de línea $ds^2 = E\cdot du^2 + 2\cdot F\cdot du\cdot dv + G\cdot dv^2$ era una propiedad intrínseca de la superficie, independiente del hecho de que estuviera embebida en el espacio euclidiano.

La geometría permanecería invariante si la superficie embebida era deformada de tal

[119] Este es un caso particular de un teorema que se verá en 9.4.1.

manera que se conservaran todas las distancias entre puntos medidas por las longitudes del arco de geodésica que los une. Diciéndolo en forma intuitiva, se conservarían las relaciones geométricas si se deformaba la superficie en el espacio sin estirarla ni cortarla. Un modelo físico de esta transformación que conserva las propiedades geométricas intrínsecas se puede construir a partir de una hoja plana de estaño muy delgada, doblándola para formar un cilindro o en general deformándola sin dilatarla ni desgarrarla. Un segmento de recta de la hoja plana, que une dos puntos próximos, se transforma en un segmento de curva geodésica que une esos mismos puntos en la superficie resultante de la deformación. Puesto que para demostrar los teoremas de la geometría euclidiana plana no es necesario hacer referencia al espacio tridimensional, todos esos teoremas pertenecen a la geometría intrínseca de cualquier superficie que se pueda obtener del plano por deformación sin estiramiento ni corte.

La relación $\alpha+\beta+\gamma-\pi=\sigma(K+\varepsilon)$ entre la suma de los ángulos de un pequeño triángulo formado por líneas geodésicas, su área σ, y la curvatura de Gauss K en un punto P de la superficie situado en el interior del triángulo (de modo que ε tiende a cero cuando el triángulo se contrae hacia el punto P y su área tiende a cero) no solo es aplicable al caso de una esfera, sino que **se aplica a cualquier superficie regular**. Reordenado esta fórmula se obtiene: $K=\dfrac{\alpha+\beta+\gamma-\pi}{\sigma}-\varepsilon$

La curvatura de Gauss es una medida de la diferencia entre la suma de los ángulos de un triángulo en el plano y en la superficie, triángulo cuyos lados en ambos casos se construyen a partir de líneas geodésicas. La Figura 9.4 muestra un triángulo de ese tipo construido en una superficie con curvatura de Gauss negativa.

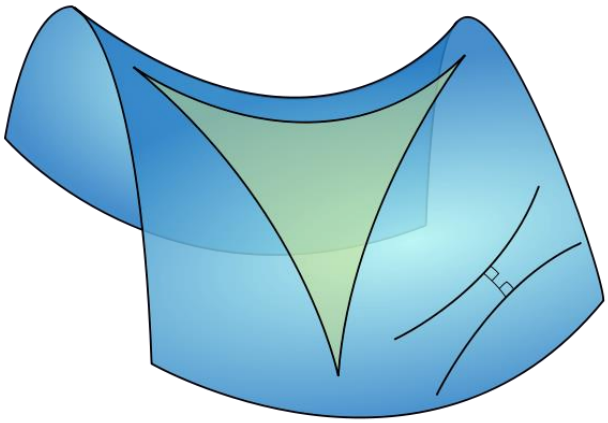


Figura 9.4 Superficie en forma de silla de montar (paraboloide hiperbólico) en la que se ha dibujado un triángulo comprendido entre líneas geodésicas. En la parte inferior derecha se puede ver el esbozo de dos arcos de curva y un arco de geodésica que corta ambos arcos en ángulo recto. En la geometría intrínseca de esa superficie son arcos de dos curvas paralelas.

Otra fórmula de la geometría intrínseca de una superficie que también presenta mucho interés, es la relación entre la longitud de una circunferencia y su radio. Ubicamos el centro de un pequeño círculo en el punto P de una superficie. Para ello definimos una circunferencia como la curva formada los puntos de la superficie que equidistan de P en este sentido: cada punto de esa curva (circunferencia) se encuentra en una línea geodésica que pasa por P y las longitudes de los correspondientes arcos de geodésica son todos iguales a un número r que por definición es el radio de la circunferencia.

Si l es la longitud de la circunferencia, se verifica: $l_c = 2\pi \cdot r - \frac{\pi}{3} \cdot K \cdot r^3 + \varepsilon_c$

En esta fórmula K es la curvatura de Gauss K en el punto P (el centro del círculo) y ε_c tiende a cero (cuando r tiende a cero) más rápido que r^3.

La curvatura de Gauss mide en este caso la diferencia entre la longitud de una circunferencia construida en la superficie a partir de arcos de geodésica de longitud r que parten de un centro P, y la expresión $2\pi \cdot r$ correspondiente a la longitud de una circunferencia en la geometría euclidiana.

De ambas fórmulas, $\alpha + \beta + \gamma - \pi = \sigma(K + \varepsilon)$ y $l = 2\pi \cdot r - \frac{\pi}{3} \cdot K \cdot r^3 + \varepsilon_c$, se desprende que cuando las dimensiones de la figura (triángulo o circunferencia) tienden a cero, **las propiedades geométricas intrínsecas a la superficie tienden a las de las figuras correspondientes de la geometría plana**.

La curvatura de Gauss permite dividir las superficies en clases de equivalencia. En cada clase se encuentran las superficies que se pueden transformar la una en la otra por deformación que conserva las propiedades métricas.

La clase con curvatura de Gauss nula incluye al plano pero no incluye esferas.

La clase con curvatura de Gauss positiva K incluye la esfera de radio $R_0 = \frac{1}{\sqrt{K}}$, puesto que en el caso de una esfera en todos los puntos los radios principales de curvatura son iguales al radio de la esfera. Pero esta clase, que incluye la esfera de radio $R_0 = \frac{1}{\sqrt{K}}$, no incluye esferas de distinto radio, porque tienen curvaturas de Gauss diferentes: las geometrías intrínsecas de esferas de distinto radio no son las mismas. Así, la suma de los ángulos de triángulos formados por geodésicas (que en el caso de la esfera son arcos de circunferencia de círculos cuyo centro es el centro de la esfera) y de áreas iguales, varía con el radio de la esfera.

Lo mismo ocurre en la clase que presenta superficies con curvatura de Gauss negativa.

Un ejemplo de superficie que presenta curvaturas de Gauss negativas es el hiperboloide de una hoja que se muestra en la Figura 9.5.

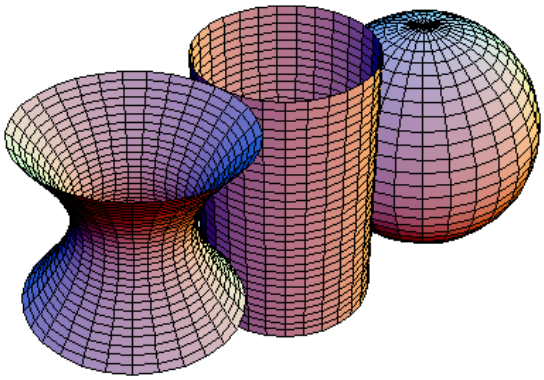


Figura 9.5 Ejemplos de superficies con curvaturas de Gauss negativas (hiperboloide de una hoja), nulas (cilindro) y positivas (esfera).

Como Gauss había llegado a la conclusión de que el postulado de las paralelas no podía ser demostrado, como sabía que era equivalente al teorema de que la suma de los ángulos de un triángulo es igual a π, y había adquirido una gran experiencia en lo concerniente a la "geometría bidimensional" de una superficie esférica, Gauss se

preguntó si nuestro espacio físico tridimensional sería realmente plano, euclidiano, o si como la superficie de una esfera, estaría curvado[120].
Eligió los picos de tres montañas en el centro de Alemania (el Inselbeg, el Brocken y el Hoher Hagen) en cuya cima colocó espejos. Supuso que un rayo de luz sería en todo caso una línea geodésica, y utilizó rayos de luz reflejados para formar un triángulo. A menos de un inevitable error experimental, concluyó que los resultados confirmaban que nuestro espacio es plano[121].
Antes de pasar a considerar, en las secciones 9.4 y 9.5, los espacios de n dimensiones y sus correspondientes geometrías, y con el propósito de comprender mejor el camino intelectual que condujo de la teoría restringida a la teoría generalizada de la relatividad, nos detendremos para analizar brevemente dos ejemplos planteados en términos de conceptos de física clásica.

9.2 ¿Curvatura o fuerza?

Resulta interesante analizar un poco más el ejemplo de una superficie esférica para ver cómo durante un movimiento restringido a esa superficie puede aparecer un fenómeno que, admitiendo la validez de la geometría euclidiana, se puede interpretar como debido a una fuerza, aunque en realidad resulta una consecuencia de la curvatura del espacio donde se produce el movimiento.
Para ello vamos a considerar un navegante en un barco, que se propone rodear una isla recorriendo los bordes de un triángulo equilátero.
Supongamos que el navegante ignora la curvatura de la Tierra, y supone la validez de la geometría euclidiana plana.
Entonces avanza, siguiendo una geodésica (que él interpreta como un segmento de recta euclidiana), una longitud igual a la del lado del triángulo, cambia de dirección haciendo un ángulo de 60º, vuelve a navegar la misma distancia, cambia nuevamente de dirección 60º, y cuando ejecuta el tramo final, se encuentra con que no termina en el punto inicial, sino en un punto interior al segmento correspondiente al primer lado del triángulo que intentaba realizar.
Todo ocurre como si la isla lo hubiera atraído hacia ella. El navegante podría suponer que esa fuerza de atracción actúa sobre el barco mientras éste recorre un lado del triángulo, de modo que para compensarla y así vencerla es necesario aumentar el ángulo con respecto a los 60º cada vez que cambia de dirección.
A partir de esta idea puede investigar el presunto campo de fuerzas asociado con la isla, encontrando que la fuerza aumenta proporcionalmente al área del triángulo, o sea aproximadamente con el cuadrado de la distancia.
En este caso sabemos que la interpretación que corresponde a la realidad es muy otra: basta considerar la curvatura de la superficie de la Tierra. Comenzando con la fórmula

---

[120] Este planteamiento se relaciona con uno muy anterior de Eratóstenes, que en una época en la cual la mayoría de las personas pensaban que la Tierra era plana, se planteó la posibilidad de que fuera esférica y diseñó un experimento, que resultó exitoso, para determinar su radio.

[121] Ahora sabemos que hubiera detectado una desviación de π inferior a 10^{-10}. En 1964, I. Shapiro midió el retardo en el eco del radar entre la Tierra y sondas espaciales ubicadas en Mercurio, Venus y Marte. Cuando el planeta con la sonda se encuentra del lado opuesto de la tierra respecto del Sol, el retardo debe aumentar respecto de lo que cabría esperar si el espacio fuera euclidiano. La mecánica celeste permite calcular las distancias euclidianas (las órbitas) con gran precisión. Sobre todo esto puede verse el libro de H. Stefani "*General relativity*", Cambridge University Press, Cambridge, UK, 1982.

para la suma de los ángulos de un triángulo esférico, se obtiene en el caso de un triángulo equilátero, para el ángulo entre dos lados: $\alpha = \frac{\pi}{3} + \frac{\sigma}{3}(K + \varepsilon)$

Se ve que es mayor que 60º, porque la curvatura de Gauss de una esfera $K = \frac{1}{R_0^2}$ es positiva. Este es el ángulo que el navegante debería haber establecido entre las navegaciones a lo largo de tramos de geodésicas de igual longitud, para que al final el triángulo se cerrara en el punto inicial. La aparente atracción de la isla es un efecto de la geometría no euclidiana bidimensional que se aplica en este caso a la superficie de una esfera de radio R_0.

9.3 Movimiento libre de una partícula newtoniana en una superficie

Supongamos ahora que una partícula se mueve en una superficie regular trazando una curva en esa superficie. Si el movimiento es por lo demás libre, la única fuerza que actúa sobre la partícula la ejerce la superficie para mantenerla allí.

Entonces en la expresión de la segunda ley de Newton $m \cdot \vec{a} = \vec{F}$, la fuerza de restricción debe ser en cada punto perpendicular a la superficie: $\vec{F} = F \cdot \vec{n}_\Sigma$ (aquí $\vec{n}_\Sigma$ es el vector unitario normal a la superficie en el punto considerado).

Por su parte, la aceleración $\vec{a}$ de la partícula se puede descomponer en una componente según el vector unitario $\vec{t}_\Gamma$ tangente a la curva Γ sobre la que se produce el movimiento (componente tangencial de la aceleración) y una componente según el vector unitario $\vec{n}_\Gamma$ conocido como normal principal a la mencionada curva (componente normal de la aceleración)[122]: $\vec{a} = \frac{dv}{dt} \cdot \vec{t}_\Gamma + v^2 \cdot \kappa \cdot \vec{n}_\Gamma$

En esta fórmula v es la magnitud de la velocidad de la partícula y κ es la curvatura de flexión de la curva.

Como consecuencia de esta descomposición y de la dirección de la fuerza de restricción la segunda ley se puede escribir así: $m \cdot \frac{dv}{dt} \cdot \vec{t}_\Gamma + m \cdot v^2 \cdot \kappa \cdot \vec{n}_\Gamma = F \cdot \vec{n}_\Sigma$

Como el vector tangente a la curva se halla en el plano tangente a la superficie en el punto considerado, se tiene que el producto escalar $\vec{t}_\Gamma \bullet \vec{n}_\Sigma = 0$

Multiplicando escalarmente ambos miembros de la expresión de la segunda ley por $\vec{n}_\Sigma$ resulta: $m \cdot v^2 \cdot \kappa \cdot (\vec{n}_\Gamma \bullet \vec{n}_\Sigma) = F$

Pero como el vector tangente y el vector normal principal son perpendiculares en un mismo punto de la curva, multiplicando ambos miembros de la fórmula de la segunda ley por $\vec{n}_\Gamma$ se obtiene: $m \cdot v^2 \cdot \kappa = F \cdot (\vec{n}_\Gamma \bullet \vec{n}_\Sigma)$

De estas dos últimas ecuaciones se desprende que siempre que $\kappa \neq 0$ entonces $(\vec{n}_\Gamma \bullet \vec{n}_\Sigma)^2 = 1$ Esta igualdad implica $(\vec{n}_\Gamma \bullet \vec{n}_\Sigma) = \pm 1$

Puesto que se trata de vectores unitarios, si suponemos que $\vec{n}_\Gamma$ apunta hacia el lado cóncavo de la curva (y de la superficie) mientras que $\vec{n}_\Sigma$ apunta en dirección opuesta,

[122] Ver, por ejemplo, el de J. Röederer "*Mecánica elemental*", 2ª edición, EUDEBA, Buenos Aires, 2002.

resulta que en cada punto de la curva trazada por la partícula en la superficie: $\vec{n}_\Gamma = -\vec{n}_\Sigma$ En caso contrario, $\vec{n}_\Gamma = \vec{n}_\Sigma$ Si se da uno de estos casos en un punto de la curva, se da en todos los puntos lo bastante próximos.

Como consecuencia de esta igualdad se deduce que $m \cdot \frac{dv}{dt} \cdot \vec{t}_\Gamma - m \cdot v^2 \cdot \kappa \cdot \vec{n}_\Sigma = F \cdot \vec{n}_\Sigma$ lo cual implica a su vez que $\frac{dv}{dt} = 0$ en todos los puntos de la curva: la magnitud de la velocidad se mantiene constante durante el movimiento.

Por otra parte, resulta que la única componente de la aceleración que en general no se anula, la aceleración normal $v^2 \cdot \kappa$, es el producto de una constante del movimiento (v^2) por la curvatura en el punto considerado.

Entonces depende solamente de una propiedad geométrica de la curva que traza la partícula en la superficie durante su movimiento: su curvatura de flexión.

Más aún, la igualdad $\vec{n}_\Gamma = \pm\vec{n}_\Sigma$ en cada punto de la curva tiene otra consecuencia notable: la curva en cuestión es una geodésica de la superficie.

Para hallar el arco de longitud mínima o en general extrema (el arco de geodésica) que une dos puntos de una superficie se aplica el cálculo de variaciones a la expresión de la longitud de un arco de curva que une esos dos puntos y está restringido a permanecer en esa superficie.

Se obtiene un sistema de ecuaciones diferenciales ordinarias[123] cuyas soluciones son las curvas geodésicas correspondientes a la superficie considerada.

Se demuestra que esas ecuaciones equivalen a la propiedad geométrica siguiente: los planos osculadores (planos que contienen el vector tangente y el vector normal principal a una curva) en cada punto de una geodésica son ortogonales al plano tangente a la superficie en ese mismo punto (es decir, el vector normal a la superficie en un punto está contenido en el plano osculador que pasa por ese mismo punto).

Como esta propiedad geométrica a su vez equivale a la igualdad $\vec{n}_\Gamma = \pm\vec{n}_\Sigma$ que obtuvimos a partir del estudio del movimiento, entonces una curva en cuyos puntos se verifique esta igualdad es una geodésica. Por ende depende solo de la superficie y es independiente de la masa de la partícula.

Ahora bien $\frac{F}{m} = -v^2 \cdot \kappa$. La magnitud de la velocidad viene determinada solo por su magnitud inicial, y la curvatura κ depende solo del punto de la superficie considerado, puesto que es la curvatura de la línea geodésica que pasa por ese punto. Por tanto el cociente $\frac{F}{m}$ no depende de la masa de la partícula: la fuerza de restricción es siempre proporcional a la masa y la aceleración en cada punto de la superficie, en forma análoga a lo que ocurre con el campo gravitatorio, es independiente de la masa de la partícula cuyo movimiento se considera.

Este ejemplo y sus generalizaciones, bien conocidos por los matemáticos y por la mayoría de los físicos a partir del último cuarto del siglo XIX, junto con la exitosa

[123] En el capítulo 7 del volumen 2 del libro de R. Courant "*Differential and integral calculus*", Interscience, New York, 1961, se puede hallar una introducción al cálculo de variaciones junto con su aplicación a la determinación de las geodésicas en una superficie en el espacio euclidiano y la interpretación geométrica de las ecuaciones diferenciales obtenidas.

descripción del movimiento en el espacio tetradimensional de Minkowski que se lleva a cabo en la teoría restringida de la relatividad, sugieren la posibilidad de abordar los fenómenos de la gravitación desde la perspectiva de un movimiento libre en un espacio abstracto curvo. En este caso cabe esperar que el tiempo, medido con respecto a un marco de referencia, ya no sea un parámetro, como en la mecánica clásica, sino que se integre junto con las variables espaciales en las coordenadas que representan los puntos de ese espacio abstracto.

9.4 Espacios abstractos y variedades de Riemann

Para formular la teoría generalizada de la relatividad es necesario, entonces, recurrir a los espacios abstractos. Riemann, en su famosa disertación para la habilitación como Profesor, que menciona Bell en la cita que da comienzo de este capítulo, mostró cómo estos espacios permiten generalizar la teoría de Gauss a continuos de más de dos dimensiones. Posteriormente, algunas de las ideas que Riemann planteó se desarrollaron dando origen al cálculo tensorial, particularmente en los trabajos de Ricci y de Levi-Cività. En esta parte se resumen algunos elementos de la teoría de espacios abstractos con el propósito de facilitar la presentación esquemática de los campos de tensores definidos sobre variedades y de la curvatura de esas variedades, que se considera en la parte 9.5.
En conjunto ambas partes están destinadas a suministrar un mínimo de herramientas matemáticas para poder discutir los fundamentos de la teoría einsteniana de la gravitación en el marco del presente trabajo.
La presentación no es rigurosa desde el punto de vista matemático (ni pretende serlo), pero se sitúa en un nivel de abstracción más elevado que el de las partes precedentes de este mismo capítulo.

9.4.1 Resumen del aparato analítico de la geometría diferencial clásica.

Esta sub-sección intenta resumir algunas de las ideas de Gauss y otros autores de trabajos matemáticos también clásicos del siglo XIX, acerca de la teoría de superficies, en notación matemática y conceptos un poco más cercanos a los actuales.
Introduciendo un sistema de coordenadas cartesianas ortogonales con origen en un punto de E_3, los puntos del espacio se pueden representar mediante vectores de posición $\vec{r} = x \cdot \vec{i} + y \cdot \vec{j} + z \cdot \vec{k}$, donde $\vec{i}, \vec{j}, \vec{k}$ los son vectores unitarios ubicados sobre los ejes de coordenadas.
El producto escalar de dos vectores $\vec{a} = a_x \cdot \vec{i} + a_y \cdot \vec{j} + a_z \cdot \vec{k}$ y $\vec{b} = b_x \cdot \vec{i} + b_y \cdot \vec{j} + b_z \cdot \vec{k}$ puede escribirse $\vec{a} \bullet \vec{b} = a_x \cdot b_x + a_y \cdot b_y + a_z \cdot b_z$
La longitud $|\vec{a}|$ de un vector se puede definir a partir del producto escalar: $|\vec{a}| = \sqrt{\vec{a} \bullet \vec{a}}$
Si ψ es en ángulo, no mayor que π, entre los vectores $\vec{a}, \vec{b}$: $\vec{a} \bullet \vec{b} = |\vec{a}| \cdot |\vec{b}| \cdot \cos\psi$
La función vectorial regular de una variable $\vec{r}(t) = x(t) \cdot \vec{i} + y(t) \cdot \vec{j} + z(t) \cdot \vec{k}$ describe una curva Γ en E_3 con vector tangente $\frac{d\vec{r}(t)}{dt} = \frac{dx(t)}{dt} \cdot \vec{i} + \frac{dy(t)}{dt} \cdot \vec{j} + \frac{dz(t)}{dt} \cdot \vec{k}$, elemento de arco $ds^2 = \frac{d\vec{r}(t)}{dt} \bullet \frac{d\vec{r}(t)}{dt} \cdot dt^2$ y longitud del arco de curva comprendido entre los puntos A y B:

$$s(\Gamma_{A \to B}) = \int_{t_A}^{t_B} \sqrt{\frac{d\vec{r}(t)}{dt} \bullet \frac{d\vec{r}(t)}{dt}} \cdot dt$$

Cuando la longitud de arco se toma como parámetro, el vector tangente $\vec{t}_\Gamma(s)=\frac{d\vec{r}(s)}{ds}=\frac{dx(s)}{ds}\cdot\vec{i}+\frac{dy(s)}{ds}\cdot\vec{j}+\frac{dz(s)}{ds}\cdot\vec{k}$ posee longitud unidad.

Como $\vec{t}_\Gamma(s)\bullet\vec{t}_\Gamma(s)=1$ para todo s, derivando respecto de la longitud de arco se obtiene:

$$\vec{t}_\Gamma(s)\bullet\frac{d}{ds}\vec{t}_\Gamma(s)=0$$

Entonces introduciendo un vector unitario $\vec{n}_\Gamma(s)$ y un escalar (que por definición es no negativo) $\kappa(s)$ (la curvatura de flexión) se puede poner: $\frac{d}{ds}\vec{t}_\Gamma(s)=\frac{d^2\vec{r}(s)}{ds^2}=\kappa(s)\cdot\vec{n}_\Gamma(s)$

El vector $\vec{n}_\Gamma(s)$ permite definir una recta orientada, la normal principal a la curva en el punto considerado.

Una superficie Σ en E_3 se puede determinar mediante la función vectorial regular de dos variables $\vec{r}(u,v)=x(u,v)\cdot\vec{i}+y(u,v)\cdot\vec{j}+z(u,v)\cdot\vec{k}$ Si se conoce esta función, cada punto de la superficie se puede identificar mediante un par de valores (u,v) en el dominio D de la función vectorial.

Dos funciones regulares $u(t)$ y $v(t)$ determinan una curva contenida en Σ cuyo vector tangente viene dado por $\frac{d\vec{r}(t)}{dt}=\frac{\partial\vec{r}(u(t),v(t))}{\partial u}\cdot\frac{du(t)}{dt}+\frac{\partial\vec{r}(u(t),v(t))}{\partial v}\cdot\frac{dv(t)}{dt}$, se encuentra en el plano tangente a la superficie y define una recta orientada en ese plano. A la línea coordenada $v=v_0$ con parámetro u le corresponde el vector tangente $\frac{\partial\vec{r}(u,v_0)}{\partial u}$, mientras que a la línea coordenada $u=u_0$ con parámetro v le corresponde el vector tangente $\frac{\partial\vec{r}(u_0,v)}{\partial v}$. El producto vectorial $\frac{\partial\vec{r}(u,v)}{\partial u}\times\frac{\partial\vec{r}(u,v)}{\partial v}$ define la dirección y sentido de la recta normal a la superficie en el punto (u,v), mientras que $\vec{N}(u,v)=\frac{\frac{\partial\vec{r}(u,v)}{\partial u}\times\frac{\partial\vec{r}(u,v)}{\partial v}}{\left|\frac{\partial\vec{r}(u,v)}{\partial u}\times\frac{\partial\vec{r}(u,v)}{\partial v}\right|}$ es el vector unitario (versor) ortogonal al plano tangente a la superficie en ese punto.

El cuadrado de la derivada del elemento de arco $\left(\frac{ds(t)}{dt}\right)^2=\frac{d\vec{r}(t)}{dt}\bullet\frac{d\vec{r}(t)}{dt}$ se puede escribir así: $\left(\frac{ds}{dt}\right)^2=E\cdot\frac{du(t)}{dt}\frac{du(t)}{dt}\cdot+2\cdot F\cdot\frac{du(t)}{dt}\cdot\frac{dv(t)}{dt}+G\cdot\frac{dv(t)}{dt}\cdot\frac{dv(t)}{dt}$

$$E(u,v)=\frac{\partial\vec{r}(u,v)}{\partial u}\bullet\frac{\partial\vec{r}(u,v)}{\partial u}\qquad F(u,v)=\frac{\partial\vec{r}(u,v)}{\partial u}\bullet\frac{\partial\vec{r}(u,v)}{\partial v}\qquad G(u,v)=\frac{\partial\vec{r}(u,v)}{\partial v}\bullet\frac{\partial\vec{r}(u,v)}{\partial v}$$

Teniendo en cuenta que $d\vec{r}=\frac{\partial\vec{r}(u,v)}{\partial u}\cdot du+\frac{\partial\vec{r}(u,v)}{\partial v}\cdot dv$:

$$ds^2=d\vec{r}\bullet d\vec{r}=E(u,v)\cdot du^2+2\cdot F(u,v)\cdot du\cdot dv+G(u,v)\cdot dv^2$$

Esta es la denominada **primera forma diferencial de la teoría de superficies**.

La longitud de un arco de curva contenido en una superficie y que une dos puntos A y B de la misma, se puede hallar empleando la primera forma diferencial:

$$s(\gamma_{A\to B})=\int_{t_A}^{t_B}\sqrt{E\cdot\frac{du(t)}{dt}\frac{du(t)}{dt}\cdot+2\cdot F\cdot\frac{du(t)}{dt}\cdot\frac{dv(t)}{dt}+G\cdot\frac{dv(t)}{dt}\cdot\frac{dv(t)}{dt}}\cdot dt$$

A partir de esta ecuación se puede plantear y resolver el problema de hallar las curvas de longitud mínima, es decir las curvas geodésicas contenidas en la superficie, como se verá en detalle en 9.5.

La segunda forma diferencial se introduce durante el estudio de cómo la superficie se encuentra inmersa en la región del espacio E_3 que la rodea, es decir, el estudio de su curvatura.

La desviación de la superficie respecto de su plano tangente en un punto (u_0,v_0) se puede caracterizar mediante la proyección del vector $\vec{r}(u,v)-\vec{r}(u_0,v_0)$ sobre el versor normal $\vec{N}(u_0,v_0)$: $\quad h(u,v)=(\vec{r}(u,v)-\vec{r}(u_0,v_0))\bullet\vec{N}(u_0,v_0)$ (Figura 9.6)

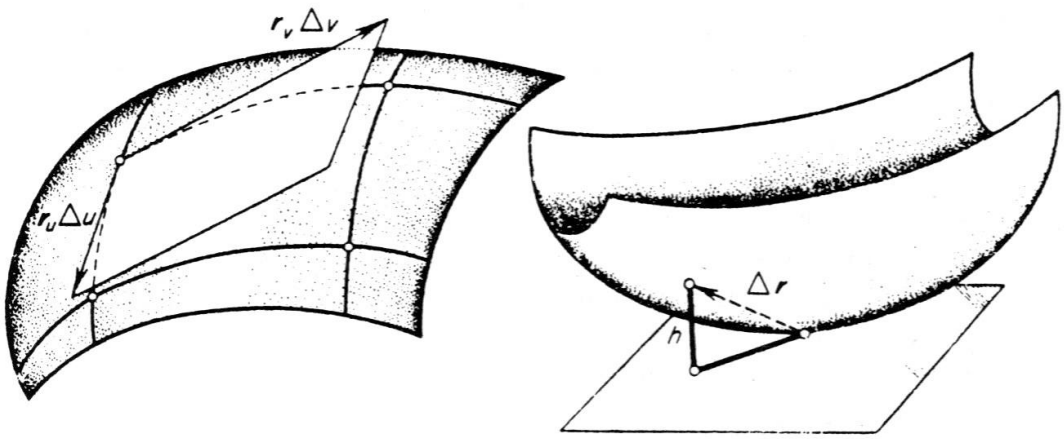


Figura 9.6

Pero si ponemos $\Delta u=u-u_0$ y $\Delta v=v-v_0$ el incremento vectorial se puede expresar mediante un desarrollo de Taylor en un entorno del punto (u_0,v_0):

$$\vec{r}(u,v)-\vec{r}(u_0,v_0)=d\vec{r}(u_0,v_0;\Delta u,\Delta v)+\frac{1}{2}\cdot d^2\vec{r}(u_0,v_0;\Delta u,\Delta v)+\vec{o}(\Delta u^2+\Delta v^2)$$

Como consecuencia la distancia entre un punto de la superficie y el plano tangente en (u_0,v_0) se puede formular así, teniendo en cuenta que $\vec{N}(u_0,v_0)\bullet d\vec{r}(u_0,v_0;\Delta u,\Delta v)=0$:

$$h(u,v)=\frac{1}{2}\cdot\vec{N}(u_0,v_0)\bullet d^2\vec{r}(u_0,v_0;\Delta u,\Delta v)+o(\Delta u^2+\Delta v^2)$$

Introduciendo los coeficientes $\quad e(u_0,v_0)=\vec{N}(u_0,v_0)\bullet\frac{\partial^2\vec{r}}{\partial u^2}(u_0,v_0)$

$$f(u_0,v_0)=\vec{N}(u_0,v_0)\bullet\frac{\partial^2\vec{r}}{\partial u\partial v}(u_0,v_0)\qquad g(u_0,v_0)=\vec{N}(u_0,v_0)\bullet\frac{\partial^2\vec{r}}{\partial v^2}(u_0,v_0)$$

$$2\cdot h(u,v)=e(u_0,v_0)\cdot\Delta u^2+2\cdot f(u_0,v_0)\cdot\Delta u\cdot\Delta v+g(u_0,v_0)\cdot\Delta v^2+o(\Delta u^2+\Delta v^2)$$

La curvatura de la sección normal de la superficie en la dirección determinada por los incrementos $\Delta u=u-u_0$ y $\Delta v=v-v_0$ cuando se fija el cociente $\lambda=\frac{\Delta v}{\Delta u}$, se puede

calcular como $k_N = \lim_{l \to 0} \frac{2 \cdot h}{l}$ donde l es la distancia en el plano tangente entre la proyección del punto de coordenadas (u, v) y el punto de coordenadas (u_0, v_0).

Resulta: $$k_N(\lambda) = e(u_0, v_0) \cdot \left(\frac{du}{ds}\right)^2 + 2 \cdot f(u_0, v_0) \cdot \left(\frac{du}{ds}\right) \cdot \left(\frac{dv}{ds}\right) + g(u_0, v_0) \cdot \left(\frac{dv}{ds}\right)^2$$

Teniendo en cuenta la primera forma fundamental esta relación se puede re-escribir:

$$k_N = \frac{e(u_0, v_0) \cdot du^2 + 2 \cdot f(u_0, v_0) \cdot du \cdot dv + g(u_0, v_0) \cdot dv^2}{E(u_0, v_0) \cdot du^2 + 2 \cdot F(u_0, v_0) \cdot du \cdot dv + G(u_0, v_0) \cdot dv^2}$$

La expresión $e(u_0, v_0) \cdot du^2 + 2 \cdot f(u_0, v_0) \cdot du \cdot dv + g(u_0, v_0) \cdot dv^2$ se conoce como la **segunda forma diferencial de la teoría de superficies**.

En función de la dirección λ: $$k_N(\lambda) = \frac{e(u_0, v_0) + 2 \cdot f(u_0, v_0) \cdot \lambda + g(u_0, v_0) \cdot \lambda^2}{E(u_0, v_0) + 2 \cdot F(u_0, v_0) \cdot \lambda + G(u_0, v_0) \cdot \lambda^2}$$

Calculando la derivada respecto de λ de la curvatura normal e igualándola a cero se puede hallar una ecuación cuya solución permite hallar los valores extremos k_1 y k_2 de $k_N(\lambda)$ junto con las direcciones principales en las que estos valores extremos son alcanzados.

Después de algunos cálculos[124], se obtiene finalmente esta ecuación de segundo grado:

$$\left(E \cdot G - F^2\right) \cdot k_N{}^2(\lambda) - \left(e \cdot G + g \cdot E - 2 \cdot f \cdot F\right) \cdot k_N(\lambda) + \left(e \cdot g - f^2\right) = 0$$

El producto de las raíces de esta ecuación es la curvatura de Gauss, intrínseca o total:

$$K = k_1 \cdot k_2 = \frac{e \cdot g - f^2}{E \cdot G - F^2}$$

La suma de las raíces es el doble de la curvatura media o de Germain $H = \frac{e \cdot G + g \cdot E - 2 \cdot f \cdot F}{E \cdot G - F^2}$ en el punto de la superficie considerado.

Finalmente, como preparación para el estudio de las geometrías en las variedades, consideremos la teoría de superficies desde un punto de vista algo diferente, debido fundamentalmente a Gauss.

La recta tangente orientada en cada punto de una curva regular contenida en una superficie determina en el plano tangente una dirección en ese punto. Esta recta junto con otra recta por ese mismo punto, tangente a la superficie y no paralela a la anterior determinan un plano tangente. A su vez ese plano tangente se puede considerar en sí mismo como un espacio euclidiano E_2, es decir, como una variedad lineal $T_2(u_0, v_0)$ asignada al punto (u_0, v_0) de la superficie. Como veremos a continuación es posible introducir todas las propiedades métricas que caracterizan la geometría en esa variedad utilizando la primera forma diferencial de la teoría de superficies, sin necesidad de recurrir al espacio E_3. Por supuesto, la variedad lineal $T_2(u_0, v_0)$, cuando el panorama se contempla desde E_3, no es otra cosa que el plano tangente a la superficie en el punto.

En general una dirección en una variedad lineal bidimensional $T_2(u_0, v_0)$ asignada a un punto (u_0, v_0) de una superficie se puede caracterizar mediante un par ordenado (a_u, a_v) de números, con al menos uno de ellos no nulo. Esos números se pueden interpretar

[124] Que se pueden encontrar en el libro de D. Struik "Geometría diferencial clásica", pp. 90-91 y 94-95, Aguilar, Madrid, 1966.

como las componentes de un vector $\vec{a}$ respecto de los vectores tangentes a las líneas coordenadas en ese punto.
A partir del elemento de línea $ds^2 = E(u_0,v_0)\cdot du^2 + F(u_0,v_0)\cdot(du\cdot dv + dv\cdot du) + G(u_0,v_0).dv^2$ en el punto (u_0,v_0) se puede definir una longitud $\|\vec{a}\|$ (la norma) del vector $\vec{a}$:

$$\|\vec{a}\|^2 = E(u_0,v_0)\cdot a_u{}^2 + F(u_0,v_0)\cdot(a_u\cdot a_v + a_v\cdot a_u) + G(u_0,v_0).a_v{}^2 > 0$$

Como esta desigualdad se verifica para todo $\vec{a}$ y en todo punto (u_0,v_0), debe verificarse $E(u_0,v_0).G(u_0,v_0) - F^2(u_0,v_0) > 0$ en todos los puntos de la superficie.
Añadiendo otra dirección $\vec{b}$ dada por sus componentes (b_u,b_v) se puede definir un producto escalar en la variedad lineal $T_2(u_0,v_0)$ en el punto (u_0,v_0):

$$(\vec{a}\bullet\vec{b})_\Sigma = E(u_0,v_0)\cdot a_u\cdot b_u + F(u_0,v_0)\cdot(a_u\cdot b_v + a_v\cdot b_u) + G(u_0,v_0).a_v\cdot b_v$$

En la definición del ángulo ψ entre $\vec{a}$ y $\vec{b}$, no mayor que π, se emplea la fórmula:

$$\cos\psi = \frac{(\vec{a}\bullet\vec{b})_\Sigma}{\|\vec{a}\|\cdot\|\vec{b}\|}$$

Por definición, dos direcciones $\vec{a}$ y $\vec{b}$ en el punto (u_0,v_0) son ortogonales entre sí cuando: $(\vec{a}\bullet\vec{b})_\Sigma = E(u_0,v_0)\cdot a_u\cdot b_u + F(u_0,v_0)\cdot(a_u\cdot b_v + a_v\cdot b_u) + G(u_0,v_0).a_v\cdot b_v = 0$
Debe tenerse en cuenta que $(\vec{a}\bullet\vec{b})_\Sigma = \vec{a}\bullet\vec{b}$ y $\|\vec{a}\| = |\vec{a}|$ para todo vector $\vec{a}$ y $\vec{b}$ del plano tangente, cuando se los considera como los vectores de E_3 que son.
Para ver esto, alcanza con constatar que $(\vec{a}\bullet\vec{b})_\Sigma$ se puede re-escribir así, teniendo en cuenta la expresión de E,F,G en términos de productos escalares en E_3 de las derivadas parciales de la función vectorial $\vec{r}(u,v)$ que define la superficie:

$$(\vec{a}\bullet\vec{b})_\Sigma = \left(\frac{\partial\vec{r}}{\partial u}(u_0,v_0)\cdot a_u + \frac{\partial\vec{r}}{\partial v}(u_0,v_0)\cdot a_v\right)\bullet\left(\frac{\partial\vec{r}}{\partial u}(u_0,v_0)\cdot b_u + \frac{\partial\vec{r}}{\partial v}(u_0,v_0)\cdot b_v\right) = \vec{a}\bullet\vec{b}$$

No obstante, lo importante es la posibilidad de definir una variedad lineal tangente en cada punto de la superficie (que como dijimos es el plano tangente, cuando se mira desde E_3) y un producto escalar en cada variedad lineal tangente, exclusivamente a partir de la primera forma diferencial **sin necesidad de recurrir a las propiedades del espacio euclidiano** en el cual la superficie se halla embebida.

Las direcciones que corresponden a las secciones normales principales son ortogonales en todos los puntos de la superficie.
La dirección $(1,0)$ del plano tangente en un punto de la superficie corresponde a la curva coordenada $v = cte$ y la $(0,1)$ corresponde a $u = cte$. Sustituyendo las componentes de esas direcciones en la expresión del producto escalar se obtiene para el valor de ese producto: $2\cdot F(u_0,v_0)$
Se deduce entonces que una condición necesaria y suficiente para que las direcciones de las líneas de coordenadas en un punto (u_0,v_0) sean ortogonales entre sí, es que se verifique: $F(u_0,v_0) = 0$

Consideremos un elemento de superficie delimitado por las líneas coordenadas que pasan por los puntos (u_0,v_0) $(u_0+\Delta u,v_0)$ $(u_0,v_0+\Delta v)$ $(u_0+\Delta u,v_0+\Delta v)$.
El área de este elemento se puede aproximar así:

$$\Delta A(u_0, v_0) = \Delta s_u \cdot \Delta s_v \cdot sen\psi + o(\Delta u^2 + \Delta v^2)$$

En esta fórmula $\Delta s_u = \sqrt{E(u_0, v_0)} \cdot \Delta u$ es la longitud de arco en la dirección $(\Delta u, 0)$, $\Delta s_v = \sqrt{G(u_0, v_0)} \cdot \Delta v$ es la longitud de arco en la dirección $(0, \Delta v)$ y ψ es el ángulo entre esas direcciones de las líneas de coordenadas que pasan por (u_0, v_0).

Aplicando la fórmula ya vista resulta: $\cos\psi = \dfrac{F(u_0, v_0)}{\sqrt{E(u_0, v_0) \cdot G(u_0, v_0)}}$

Calculando el seno de ψ y sustituyendo en la expresión para el elemento de área resulta:

$$\Delta A(u_0, v_0) = \sqrt{E(u_0, v_0) \cdot G(u_0, v_0) - F^2(u_0, v_0)} \cdot \Delta u \cdot \Delta v + o(\Delta u^2 + \Delta v^2)$$

La expresión $\sqrt{E(u_0, v_0) \cdot G(u_0, v_0) - F^2(u_0, v_0)} \cdot \Delta u \cdot \Delta v$ es el área de un elemento de superficie en el plano tangente en el punto (u_0, v_0). Como se estudia en detalle en los cursos de cálculo infinitesimal, sumando un conjunto apropiado de áreas como ésta se puede obtener una aproximación al área de una porción de superficie. En la medida en que el área $A(\Sigma)$ de una superficie es el límite al que esas sumas, se puede hallar como una integral doble extendida al dominio D de variación de los parámetros de la superficie:

$$A(\Sigma) = \iint_D \sqrt{E(u,v).G(u,v) - F^2(u,v)}.du \cdot dv$$

Si $ds^2 = E_0(u,v) \cdot du^2 + 2 \cdot F_0(u,v) \cdot du \cdot dv + G_0(u,v).dv^2$ es el elemento de línea sobre una esfera, el elemento de área sobre la esfera viene dado por:

$$\Delta\Omega(u_0, v_0) = \sqrt{E_0(u_0, v_0) \cdot G_0(u_0, v_0) - F_0{}^2(u_0, v_0)} \cdot \Delta u \cdot \Delta v + o(\Delta u^2 + \Delta v^2)$$

Si B es el dominio de variación de los parámetros, el área de un casquete esférico Ψ viene dada por: $\Omega(\Psi) = \iint_B \sqrt{E_0(u,v).G_0(u,v) - F_0{}^2(u,v)}.du \cdot dv$

Estos resultados se pueden aplicar a una porción de superficie y a su representación esférica vistas en 9.1.1 como parte de la definición de la curvatura de Gauss, (considerada como una medida de la divergencia de las normales a la superficie en el entorno de un punto).

Si $\vec{N}(u,v)$ representa el campo de versores normales a la porción de superficie, el elemento de línea de la representación esférica correspondiente se denomina **tercera forma fundamental** de la teoría de superficies y viene dado por $ds^2 = d\vec{N}(u,v) \bullet d\vec{N}(u,v)$

En ese caso podemos utilizar un dominio D de parámetros común al elemento de superficie y a su representación esférica. Si contraemos el dominio en torno a un punto fijo (u_*, v_*), podemos calcular el valor absoluto de la curvatura intrínseca en el correspondiente punto de la superficie como el límite:

$$|K(u_*, v_*)| = \lim_{A(\Sigma)\to 0}\left(\frac{\Omega(\Psi)}{A(\Sigma)}\right) = \frac{\sqrt{E_0(u_*, v_*).G_0(u_*, v_*) - F_0{}^2(u_*, v_*)}}{\sqrt{E(u_*, v_*).G(u_*, v_*) - F^2(u_*, v_*)}}$$

La definición de la curvatura de Gauss como producto de las curvaturas principales y la definición, que acabamos de ver, basada en una medida de la divergencia de las normales, **son equivalentes**, debido a que se verifica[125]:

[125] Esta igualdad queda demostrada si podemos probar esta otra:

$$\lim_{A(\Sigma)\to 0}\left(\frac{\Omega(\Psi)}{A(\Sigma)}\right)=\frac{\sqrt{E_0(u,v).G_0(u,v)-F_0{}^2(u,v)}}{\sqrt{E(u,v).G(u,v)-F^2(u,v)}}=\frac{\left|e(u,v)\cdot g(u,v)-f^2(u,v)\right|}{E(u,v).G(u,v)-F^2(u,v)}=\left|k_1\cdot k_2\right|$$

Es posible demostrar que los coeficientes $E_0(u,v)$ $F_0(u,v)$ $G_0(u,v)$ del elemento de línea sobre la esfera se pueden hallar a partir de las funciones $E(u,v)$ $F(u,v)$ $G(u,v)$ y de sus derivadas parciales hasta el segundo orden.

Como consecuencia, la curvatura de Gauss se puede hallar a partir del elemento de línea de la superficie, sin necesidad de información adicional.

Este resultado es teorema egregio de Gauss. En particular, cuando las líneas de coordenadas son ortogonales, la curvatura intrínseca se puede calcular en términos de los coeficientes $E(u,v)$ y $G(u,v)$ que no se anulan, mediante la fórmula siguiente[126]:

$$K=-\frac{1}{\sqrt{E\cdot G}}\left[\frac{\partial}{\partial u}\left(\frac{1}{\sqrt{E}}\cdot\frac{\partial\sqrt{G}}{\partial u}\right)+\frac{\partial}{\partial v}\left(\frac{1}{\sqrt{G}}\cdot\frac{\partial\sqrt{E}}{\partial v}\right)\right]$$

Una superficie con curvatura de Gauss variable, que inclusive cambia de signo al pasar de una porción de la superficie a otra, es geométricamente muy heterogénea porque como consecuencia de los cambios de curvatura las propiedades de su geometría intrínseca se modifican de un punto a otro. Esta heterogeneidad hace que la geometría intrínseca en una superficie pueda diferir mucho de la geometría plana, pese a ciertas similitudes de índole general, como la correspondencia entre rectas y geodésicas, y entre ángulos determinados por la intersección de rectas y ángulos determinados por geodésicas que se intersecan.

Supongamos que en una superficie formamos un triángulo delimitado por curvas geodésicas. El dominio de parámetros (u,v) del triángulo es un subdominio D_Δ del dominio D de la superficie. Mientras que en el plano la suma de los ángulos de un triángulo siempre es igual a la suma de dos ángulos rectos, en una superficie cualquiera la suma de los ángulos de un triángulo geodésico no está determinada a menos que conozcamos la curvatura de Gauss de la superficie.

La suma de los ángulos se relaciona con la curvatura de Gauss de esta manera:

$\alpha+\beta+\gamma=\pi+\iint_{D_\Delta}K(u,v)\cdot dA(u,v)$, siendo $dA(u,v)=\sqrt{E(u,v).G(u,v)-F^2(u,v)}\cdot du\cdot dv$

$$\left|e(u,v)\cdot g(u,v)-f^2(u,v)\right|=\sqrt{E_0(u,v).G_0(u,v)-F_0{}^2(u,v)}\cdot\sqrt{E(u,v).G(u,v)-F^2(u,v)}$$

Para ello se utilizan dos herramientas:

-Una forma alternativa de escribir los coeficientes de la segunda forma diferencial de la teoría de superficies que se sustituye en la expresión $e(u,v)\cdot g(u,v)-f^2(u,v)$, a saber:

$$e(u,v)=-\frac{\partial\vec{N}(u,v)}{\partial u}\bullet\frac{\partial\vec{r}}{\partial u}(u,v)\qquad f(u,v)=-\frac{\partial\vec{N}(u,v)}{\partial u}\bullet\frac{\partial\vec{r}}{\partial v}(u,v)=-\frac{\partial\vec{N}(u,v)}{\partial v}\bullet\frac{\partial\vec{r}}{\partial u}(u,v)\qquad g(u,v)=-\frac{\partial\vec{N}(u,v)}{\partial v}\bullet\frac{\partial\vec{r}}{\partial v}(u,v)$$

-La identidad de Lagrange: $(\vec{a}\times\vec{b})\bullet(\vec{c}\times\vec{d})=(\vec{a}\bullet\vec{c})\cdot(\vec{b}\bullet\vec{d})-(\vec{b}\bullet\vec{c})\cdot(\vec{a}\bullet\vec{d})$ que se utiliza para deducir que $e\cdot g-f^2=\left(\frac{\partial\vec{N}}{\partial u}\times\frac{\partial\vec{N}}{\partial v}\right)\bullet\left(\frac{\partial\vec{r}}{\partial u}\times\frac{\partial\vec{r}}{\partial u}\right)=\left(\frac{\partial\vec{N}}{\partial u}\times\frac{\partial\vec{N}}{\partial v}\right)\bullet\left(\left|\frac{\partial\vec{r}}{\partial u}\times\frac{\partial\vec{r}}{\partial u}\right|\cdot\vec{N}\right)$

La igualdad entre las dos definiciones de la curvatura intrínseca se desprende entonces de las ecuaciones:

$$\left(\frac{\partial\vec{N}}{\partial u}\times\frac{\partial\vec{N}}{\partial v}\right)\bullet\vec{N}=\pm\left|\left(\frac{\partial\vec{N}}{\partial u}\times\frac{\partial\vec{N}}{\partial v}\right)\right|=\pm\sqrt{E_0(u,v).G_0(u,v)-F_0{}^2(u,v)}\qquad\left|\frac{\partial\vec{r}}{\partial u}\times\frac{\partial\vec{r}}{\partial u}\right|=\sqrt{E(u,v).G(u,v)-F^2(u,v)}$$

[126] Ver el libro de D. Struik "Geometría diferencial clásica", pp. 127-128, Aguilar, Madrid, 1966.

Si σ es el área del triángulo y (u_0, v_0) es un punto interior al mismo, teniendo en cuenta que $K(u,v)$ es una función regular, entonces $\iint_{D_\Delta} K(u,v)\cdot dA(u,v)$ se puede aproximar por $K(u_0,v_0)\cdot\sigma + o(\sigma)$ de modo que resulta: $\alpha+\beta+\gamma-\pi=\sigma(K+\varepsilon)$

En esta expresión $o(\sigma)=\varepsilon\cdot\sigma$, y ε tiende a cero cuando σ tiende a cero. Se obtiene así la fórmula utilizada en la discusión de la geometría intrínseca en la sub-sección 9.1.2.

Las ideas fundamentales de la geometría diferencial clásica fueron extendidas por Riemann y otros matemáticos, a partir de mediados del siglo XIX, sentando las bases de una teoría de los espacios abstractos de n dimensiones que constituye la herramienta matemática básica de la relatividad generalizada. El resto del presente capítulo se destina a estudiar esa teoría generalizada.

9.4.2 Espacios euclidianos de n dimensiones

Eligiendo un origen O y un sistema de coordenadas cartesianas ortogonales se establece una correspondencia uno a uno entre puntos P y ternas ordenadas de números reales (x^1, x^2, x^3), o en forma equivalente entre vectores de posición $\overrightarrow{OP}$ y ternas (x^1, x^2, x^3). A cada vector $\overrightarrow{OP}$ le corresponde una terna y a cada terna un vector.

En el conjunto de ternas se puede definir una adición de ternas, el producto de una terna por un escalar (número real) y el producto interno o escalar de dos ternas:

$$(x^1, x^2, x^3)+(\tilde{x}^1, \tilde{x}^2, \tilde{x}^3)=(x^1+\tilde{x}^1, x^2+\tilde{x}^2, x^3+\tilde{x}^3)$$

$$w\cdot(x^1, x^2, x^3)=(w\cdot x^1, w\cdot x^2, w\cdot x^3)$$

$$(x^1, x^2, x^3)\bullet(\tilde{x}^1, \tilde{x}^2, \tilde{x}^3)=x^1\cdot\tilde{x}^1+x^2\cdot\tilde{x}^2+x^3\cdot\tilde{x}^3$$

A partir del producto interno $(\tilde{x}^1-x^1, \tilde{x}^2-x^2, \tilde{x}^3-x^3)\bullet(\tilde{x}^1-x^1, \tilde{x}^2-x^2, \tilde{x}^3-x^3)$ se define la distancia euclidiana entre dos puntos $\sqrt{(\tilde{x}^1-x^1)^2+(\tilde{x}^2-x^2)^2+(\tilde{x}^3-x^3)^2}$

Entonces en el conjunto de las ternas ordenadas de números reales queda definida una estructura de espacio vectorial con producto interno y métrica inducida por ese producto interno. A partir de esta estructura y de la correspondencia entre ternas y puntos, se puede desarrollar la geometría euclidiana tridimensional en forma enteramente analítica. Aunque no se la pueda visualizar mediante un modelo físico, esta representación se puede extender[127] para definir un espacio euclidiano de dimensión N representando los puntos del espacio mediante el conjunto de N-uplas ordenadas de números reales $(x^1, x^2, ..., x^N)$. Las x^j constituyen las coordenadas respecto de cierto sistema de coordenadas cartesianas ortogonales, con la distancia entre puntos definida $\sqrt{\sum_{k=1}^{N}(\tilde{x}^k-x^k)^2}$ a partir del producto interno:

$$(x^1, x^2, ..., x^N)\bullet(\tilde{x}^1, \tilde{x}^2, ..., \tilde{x}^N)=\sum_{k=1}^{N} x^k\cdot\tilde{x}^k$$

[127] Al igual que la correspondiente geometría N-dimensional.

Esta distancia se puede expresar en forma infinitesimal mediante el cuadrado del elemento de distancia entre dos puntos $\left(x^1, x^2, ..., x^N\right)$ y $\left(x^1 + dx^1, x^2 + dx^2, ..., x^N + dx^N\right)$:

$$ds^2 = \sum_{k=1}^{k=N} \left(dx^k\right)^2$$

En principio se debería poder utilizar cualquier sistema de coordenadas (no necesariamente coordenadas cartesianas ortogonales), que dejara invariante la expresión de ds^2 en cada punto del espacio.

Un cambio de coordenadas regular, que conserve ds^2 pero por lo demás arbitrario, $x'^j = f^j\left(x^1, x^2, ...x^N\right)$ para j=1, 2,…, N, en general conduce a una expresión del tipo $ds^2 = \sum_{i,k=1}^{N} g'_{i,k} dx'^i dx'^k$ donde los N^2 coeficientes dependen en general de las coordenadas $\left(x'^1, x'^2, ..., x'^N\right)$ pero se pueden tomar siempre simétricos: $g'_{i,k} = g'_{k,i}$

En consecuencia, solamente hay $\frac{N \cdot (N+1)}{2}$ coeficientes independientes.

Por ejemplo, para N=3, en coordenadas esféricas $x'^1 = r, x'^2 = \theta, x'^3 = \varphi$ se tiene $g'_{11} = 1, g'_{22} = r^2, g'_{33} = r^2 \cdot sen^2\theta$ mientras que los demás coeficientes se anulan. Entonces: $ds^2 = dr^2 + r^2 \cdot d\theta^2 + r^2 \cdot sen^2\theta \cdot d\varphi^2$

Pero lo característico de un espacio euclidiano generalizado es que siempre se pueden hallar cambios de coordenadas que permiten llevar el cuadrado de la distancia infinitesimal a la forma $ds^2 = \sum_{k=1}^{k=N} \left(dx^k\right)^2$. El mismo sistema de coordenadas (denominado sistema de coordenadas cartesianas ortogonales N-dimensional) permite expresar, para todos los puntos del espacio E_N, el elemento de línea de esta forma particular, denominada forma diagonal.

Como un espacio euclidiano generalizado no está en principio referido a ningún otro, la geometría que se desarrolle en él puede considerarse una geometría intrínseca.

En el caso de las superficies inmersas en el espacio euclidiano tridimensional, vimos en 9.1.2 que es posible definir una geometría intrínseca para una superficie a partir del elemento de línea $ds^2 = E \cdot du^2 + 2 \cdot F \cdot du \cdot dv + G \cdot dv^2$.

También vimos que la propiedad más notable de la curvatura de Gauss es que las superficies con diferentes curvaturas de Gauss son inherentemente distintas entre sí: si las deformamos de todas las maneras posibles, sin estirarlas ni romperlas, nunca podremos superponer una sobre otra.

Entonces resulta natural generalizar estos resultados a espacios abstractos, denominados variedades, que son el equivalente n-dimensional de las superficies bidimensionales del espacio euclidiano de la geometría diferencial clásica.

9.4.3 Variedades riemannianas de n dimensiones

Desde una perspectiva intuitiva se puede decir que una **variedad n-dimensional** es un conjunto de objetos P, denominados puntos, relacionados entre sí de tal forma que se asemeja localmente a una porción deformada de un espacio euclidiano n-dimensional[128]. Esta semejanza permite representar los puntos de al menos una parte de la variedad mediante una n-upla de coordenadas pertenecientes a un espacio euclidiano.

No obstante esta semejanza local, la estructura global de una variedad puede ser muy diferente de la de un espacio euclidiano.

Por ejemplo, una superficie esférica, la superficie de un toro (un modelo matemático de una rosquilla) y la superficie de un cilindro infinito de sección circular son variedades bidimensionales formadas por puntos del espacio euclidiano tridimensional E_3.

Si bien localmente, en la vecindad de cada uno de sus puntos, todas ellas se asemejan a una porción del plano euclidiano, vistas globalmente son muy diferentes.

Una diferencia importante entre esas variedades es la siguiente: cualquier curva cerrada contenida en la superficie de una esfera se puede contraer hacia un punto sin abandonar la variedad, mientras que eso no es posible para ciertas curvas definidas en la superficie del toro y en la del cilindro.

Una **variedad de Riemann**[129] n-dimensional V_n es un conjunto de puntos representables mediante n-uplas ordenadas de números reales $\left(x^1, x^2, ..., x^n\right)$ tal que en

---

[128] Las variedades son un tipo especial de espacio topológico. A cada punto P de la variedad le corresponde en forma biunívoca (es decir, uno a uno) una n-upla $\left(x^1, x^2, ..., x^n\right)$ de números reales, cuyos componentes se denominan coordenadas del punto P. A la correspondencia P$\leftrightarrow\left(x^1, x^2, ..., x^n\right)$ entre puntos y n-uplas se la denomina sistema de coordenadas (o carta) definido en la variedad. En general se precisa más de una carta para representar la totalidad de los puntos de una variedad. Un conjunto de cartas adecuado forma un atlas de coordenadas. Si P$\leftrightarrow\left(x'^1, x'^2, ..., x'^n\right)$ es otro sistema de coordenadas, entonces hay una relación uno a uno entre las coordenadas de cada punto así representado: $x'^k = f_k\left(x^1, x^2, ..., x^n\right)$ $x^l = f_l^{-1}\left(x'^1, x'^2, ..., x'^n\right)$ $k, l = 1,2,...,n$ La n-upla formada por las funciones f_l^{-1} constituye la función inversa de la n-upla formada por las funciones f_k. Supondremos que estas funciones son regulares (poseen derivadas parciales continuas de todos los órdenes). La variedad correspondiente se denomina variedad regular. Se pueden hallar definiciones concisas de estos y otros conceptos relacionados con las variedades (en inglés, manifolds) en el libro editado por Ema Previato "*Dictionary of applied mathematics for engineers and scientists*", CRC press, Boca Raton, USA, 2003.

[129] Los contenidos de cálculo tensorial y geometrías de Riemann necesarios para comprender los fundamentos matemáticos, tanto la teoría de la gravitación de Einstein como sus principales aplicaciones cosmológicas, se pueden estudiar en el notable y didáctico libro de R. Lambourne "*Relativity, gravitation and cosmology*", co-editado por The Open University y por Cambridge University Press, Cambridge, U.K., 2010. Entre las muchas obras sobre tensores y sus aplicaciones a las geometrías de Riemann se tienen a nivel intermedio el libro de L. Santaló "*Vectores y tensores con sus aplicaciones*", EUDEBA, Buenos Aires, 1969, y el libro de A. Mc Connell "*Applications of tensor analysis*", Dover, New York, 1947. A nivel avanzado se tiene los libros de H. Weyl "*Space, time, matter*", Dover, New York, 1950; de J.L.Synge "*Relativity: the general theory*", North-Holland, Amsterdam, 1960; de T. Thomas "*Tensor analysis and differential geometry*", Academic Press, New York, 1965; de D.Lovelock y H.Rund "*Tensors, differential forms and variational principles*" Wiley, New York, 1975; y de S.Hasani "*Mathematical Physics*" Springer, New York, 2000.

cada punto se encuentra definido una forma cuadrática en términos de los diferenciales de las coordenadas: $ds^2 = \sum_{i,k=1}^{n} g_{i,k} dx^i dx^k$

Esta forma debe permanecer invariante bajo los cambios de coordenadas que se puedan aplicar en la variedad.

Riemann la interpretaba como una expresión del cuadrado de la distancia entre puntos de coordenadas $(x^1, x^2, ..., x^n)$ y $(x^1 + dx^1, x^2 + dx^2, ..., x^n + dx^n)$ "infinitamente próximos", como se decía a mediados del siglo XIX.

Los coeficientes $g_{i,k}$ son funciones regulares de las coordenadas $(x^1, x^2, ..., x^n)$ con excepción, quizás, de algunos puntos donde presentan singularidades[130].

Suponemos que $\sum_{i,k=1}^{n} g_{i,k} dx^i dx^k$ es o bien definida positiva (es decir es positiva excepto cuando todos los dx^i son nulos, en cuyo caso se anula) o si no lo es, por lo menos el determinante de la matriz de coeficientes $g(x^1, x^2, ..., x^n) = \det[g_{i,k}(x^1, x^2, ..., x^n)] \neq 0$ en todo punto de la variedad.

Si la forma diferencial $ds^2 = \sum_{i,k=1}^{n} g_{i,k} dx^i dx^k$ es definida positiva, define una métrica en la variedad como la concebía Riemann.

Si la forma diferencial $ds^2 = \sum_{i,k=1}^{n} g_{i,k} dx^i dx^k$ no es definida positiva, entonces cambia de signo al variar los incrementos en las coordenadas y define una pseudo-métrica. Este último caso es el que interesa en el caso de la teoría de la relatividad.

El espacio euclidiano (si el elemento de línea ds^2 es definido positivo) o pseudo-euclidiano (si ds^2 no es definido positivo) de n dimensiones, son los casos particulares más simples de una variedad n-dimensional. Se las denomina, respectivamente, **variedades lineales euclidiana** o **pseudo-euclidiana**.

9.4.4 Variedades lineales tangentes

Además, las propiedades geométricas de una variedad en la vecindad de un punto de V_n se aproximan tanto más a las de una variedad euclidiana o pseudo-euclidiana del mismo número de dimensiones cuanto menor sea el radio de esa vecindad. Se tiene así el denominado **espacio tangente (o variedad lineal tangente)** a la variedad de Riemann en ese punto. Se puede considerar como una generalización natural del concepto de plano tangente a una superficie bidimensional.

Si $(x^1, x^2, ..., x^n)$ es un punto de V_n , el espacio tangente a la variedad en ese punto es la variedad lineal (euclidiana o pseudoeuclidiana) $E_n(x^1, x^2, ..., x^n)$ de n dimensiones cuyos puntos se representan mediante $(y^1, y^2, ..., y^n)$ (de modo que el punto $(0,0,...,0)$ de

---

[130] Los coeficientes $g_{i,k}$ pueden no estar acotados en un entorno de uno de esos puntos, como ocurre en el caso de la métrica de Schwarzschild que se estudiará en el capítulo 11 y que constituye una de las pocas soluciones analíticas exactas que se conocen de las ecuaciones de la teoría de la gravitación de Einstein.

la variedad tangente corresponde al punto $\left(x^1, x^2, ..., x^n\right)$ de V_n) y que posee la forma diferencial cuadrática asociada: $ds_{\tan}{}^2 = \sum_{i,k=1}^{n} g_{i,k}\left(x^1, x^2, ..., x^n\right) \cdot dy^i dy^k$

Esta forma define ya sea una métrica euclidiana, ya sea una pseudométrica euclidiana.

Los coeficientes $g_{i,k}\left(x^1, x^2, ..., x^n\right)$ de una forma cuadrática se pueden considerar simétricos, de modo que se puede hallar un cambio lineal de coordenadas[131] en la variedad lineal tangente tal que en las nuevas coordenadas $\left(z^1, z^2, ..., z^n\right)$ el elemento de línea se reduce a la forma diagonal $ds^2 = \sum_{l=1}^{n} \varepsilon_l \cdot dz_l^2$ donde cada ε_l vale ± 1. El número de coeficientes positivos p dentro del total de n coeficientes (en este caso el resto $n-p$ de los ε_l son negativos) define **la signatura** $\sigma = p - (n-p) = 2 \cdot p - n$ del elemento de línea en tanto forma cuadrática.

Como los coeficientes $g_{i,k}\left(x^1, x^2, ..., x^n\right)$ son funciones regulares de las coordenadas $\left(x^1, x^2, ..., x^n\right)$ del punto de la variedad V_n considerado, las signaturas de las formas cuadráticas son las mismas[132] en cada uno de los puntos de la variedad.

A partir de la forma cuadrática $\sum_{i,k=1}^{n} g_{i,k}\left(x^1, x^2, ..., x^n\right) \cdot y^i \cdot y^k$ se puede construir una forma bilineal $\sum_{i,k=1}^{n} g_{i,k}\left(x^1, x^2, ..., x^n\right) \cdot y^i \cdot z^k$ en la variedad lineal tangente.

Introduciendo una base $\left(\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\right)$ se puede introducir un pseudo-producto escalar de dos vectores $\vec{y} = \sum_{i=1}^{n} y^i \cdot \vec{e}_i$ y $\vec{z} = \sum_{k=1}^{n} z^k \cdot \vec{e}_k$ definiendo $\vec{e}_i \bullet \vec{e}_k = g_{i,k}$ para $i, k = 1,2,...,n$

Se tiene así: $\vec{y} \bullet \vec{z} = \left(\sum_{i=1}^{n} y^i \cdot \vec{e}_i\right) \bullet \left(\sum_{k=1}^{n} z^k \cdot \vec{e}_k\right) = \sum_{i,k=1}^{n} \vec{e}_i \bullet \vec{e}_k \cdot y^i z^k = \sum_{i,k=1}^{n} g_{i,k} \cdot y^i z^k$

A diferencia de un verdadero producto escalar, que verifica $\vec{y} \bullet \vec{y} > 0$ para todo $\vec{y} \neq \vec{0}$ y se anula para $\vec{y} = \vec{0}$, el pseudo-producto escalar no posee un signo definido y para algunos vectores se verifica $\vec{y} \bullet \vec{y} = 0$ siendo $\vec{y} \neq \vec{0}$. No obstante comparte con el producto escalar las demás propiedades básicas.

---

[131] Esto alude al bien conocido proceso de diagonalización de una forma cuadrática mediante transformaciones lineales de coordenadas. Las formas bilineales y los espacios pseudo-euclidianos se consideran en detalle en el capítulo 6 del libro de Máltsev "*Fundamentos de álgebra lineal*", Mir, Moscú, 1972.

[132] En el caso de la relatividad generalizada, la variedad de Riemann posee cuatro dimensiones. Además, como se verá en el próximo capítulo, se supone que esta variedad posee un elemento de línea que generaliza el intervalo del tiempo-espacio de Minkowski $ds^2 = c^2(dt)^2 - (dx)^2 - (dy)^2 - (dz)^2$ de tal forma que la variedad lineal tangente en cada punto es un espacio pseudo-euclidiano de Minkowski. Entonces la signatura del elemento de línea de la variedad de Riemann es la misma que la del intervalo de tiempo- espacio.

Por ejemplo, en el espacio lineal pseudo-euclidiano de Minkowski, a cada marco de referencia inercial le corresponde una base[133] $(\vec{e}_0, \vec{e}_1, \vec{e}_2, \vec{e}_3)$ tal que $\vec{e}_0 \bullet \vec{e}_0 = 1$, $\vec{e}_j \bullet \vec{e}_j = -1$ para $j = 1,2,3$ mientras que $\vec{e}_i \bullet \vec{e}_k = 0$ siempre que $i \neq k$.

Entonces para un cuadrivector $\vec{y} = \sum_{i=0}^{3} y^i \cdot \vec{e}_i$ de ese espacio se tiene:

$$\vec{y} \bullet \vec{y} = (y^0)^2 - (y^1)^2 - (y^2)^2 - (y^3)^2$$

Si $y^0 = c \cdot t$, y si las coordenadas y^1, y^2, y^3 representan puntos del espacio físico (tridimensional euclidiano) donde se halla el frente de onda correspondiente a un destello luminoso producido en el origen en el instante inicial $t = 0$ y que se propaga en el vacío, entonces para ese cuadrivector $\vec{y} \bullet \vec{y} = 0$ aunque $\vec{y} \neq \vec{0}$.

Cuando se tiene una métrica propiamente dicha, se pueden definir conceptos geométricos que extienden la teoría de superficies a n dimensiones, puesto que se puede definir productos internos en todas las variedades lineales tangentes.
Así se puede introducir el ángulo entre dos direcciones $\vec{a}$ y $\vec{b}$ dadas en la variedad lineal $T_n(P)$ tangente a un punto P de V_n: $\cos\psi = \dfrac{\vec{a} \bullet \vec{b}}{\|\vec{a}\| \cdot \|\vec{b}\|}$ $\quad \|\vec{a}\|^2 = \vec{a} \bullet \vec{a} \quad \|\vec{b}\|^2 = \vec{b} \bullet \vec{b}$

Esto a su vez permite extender la definición de direcciones ortogonales (vista en 9.1.2 para superficies bidimensionales) a cada punto de la variedad.
Una tarea importante es buscar una expresión analítica de la diferencia local, es decir, la diferencia en las proximidades de un punto, entre una variedad V_n y el espacio euclidiano o pseudo-euclidiano que le es tangente en ese punto. Esa expresión analítica caracteriza la curvatura de la variedad en un punto dado.
Es importante remarcar que el concepto de curvatura de una variedad, en el sentido de su geometría intrínseca, no tiene en principio nada que ver con la idea de que la variedad se encuentra inmersa en un espacio mayor, que la rodea y en el cual la variedad de algún modo se curva, como vemos que se curva en el espacio físico de nuestra percepción, un modelo de superficie hecho mediante una hoja delgada de estaño. La curvatura intrínseca se define a partir de las propiedades métricas propias de la variedad considerada, y mide la diferencia entre las propiedades geométricas locales y las propiedades que posee un espacio euclidiano o pseudo-euclidiano de igual dimensión, la variedad lineal tangente. Será considerada en 9.5.4.

9.4.5 Inmersión de variedades riemannianas en espacios euclidianos

Pero también es lícito preguntarse por la geometría extrínseca, la geometría que depende de cómo la variedad se halla inmersa en un espacio euclidiano generalizado. Esto a su vez conduce a preguntarse por la dimensión N de un espacio euclidiano E_N, espacio en el cual la variedad n-dimensional pueda ser embebida.
Supongamos que las coordenadas cartesianas ortogonales de los puntos de E_N son $(X^1, X^2, ..., X^N)$. Entonces V_n se puede caracterizar por N funciones $X^k = F^k(x^1, x^2, ..., x^n)$ k=1,2,…, N

[133] En el caso del espacio de Minkowski numeramos los vectores de la base y las coordenadas a partir de 0 en lugar de 1 para destacar el rol especial del tiempo entre las cuatro coordenadas de un evento.

Mientras que las coordenadas $\left(x^1, x^2, ..., x^n\right)$ son intrínsecas a la variedad, como lo son las coordenadas (u, v) empleadas por Gauss en su teoría de superficies, las funciones $X^k = F^k\left(x^1, x^2, ..., x^n\right)$ dan las coordenadas cartesianas extrínsecas a la variedad, al igual que las funciones $x = x(u, v) \quad y = y(u, v) \quad z = z(u, v)$ dan las coordenadas cartesianas extrínsecas de una superficie en la teoría de Gauss.

El elemento de línea en E_N, es decir $ds^2 = \sum_{k=1}^{k=N}\left(dX^k\right)^2$, se debe reducir al elemento de línea sobre la variedad $ds^2 = \sum_{i,k=1}^{n} g_{i,k} dx^i dx^k$ cuando las coordenadas en E_N se expresan mediante las funciones $X^k = F^k\left(x^1, x^2, ..., x^n\right)$.

Para que en E_N se puedan embeber todas las variedades de dimensión n, el número N de funciones de $\left(x^1, x^2, ..., x^n\right)$ debe ser suficiente para determinar los coeficientes $g_{i,k}\left(x^1, x^2, ..., x^n\right)$ para todas las variedades posibles.

Como por simetría solo hay $\frac{n \cdot (n+1)}{2}$ coeficientes independientes, el número mínimo de funciones $X^k = F^k\left(x^1, x^2, ..., x^n\right)$ y por tanto la mínima dimensión N del espacio euclidiano generalizado debe verificar: $N = \frac{n \cdot (n+1)}{2}$

Si n=2 (caso de las superficies estudiadas por Gauss), resulta N=3. Pero si **n=4** (caso del tiempo-espacio curvo de la teoría general de la relatividad) resulta **N=10**.

Si la variedad V_n se embebe en un espacio euclidiano, por cada punto de la variedad se puede construir una variedad lineal tangente n-dimensional que cumple el mismo rol geométrico que el plano tangente a una superficie del espacio tridimensional.

Constituye una aproximación lineal a la variedad en un entorno del punto.

Igual que en el caso de las superficies bidimensionales, se puede analizar el apartamiento (las curvaturas) de la variedad respecto su aproximación lineal en cada punto. Todo esto es parte de la geometría extrínseca.

Consideremos ahora algunos ejemplos de variedades[134] embebidas en un espacio euclidiano de 4 dimensiones, E_4, de coordenadas cartesianas ortogonales x_1, x_2, x_3, x_4.

La ecuación $x_1^2 + x_2^2 + x_3^2 + x_4^2 = R^2$ define una variedad esférica de 3 dimensiones (**3-esfera**), de radio R. Si anulamos una de las coordenadas cartesianas, se obtiene como resultado que la traza de la esfera en el correspondiente sub-espacio euclidiano tridimensional es otra esfera, ahora bidimensional, con el mismo radio.

[134] Como se verá en el capítulo 12 sobre el problema cosmológico, algunas de estas variedades tridimensionales aparecen como modelos en gran escala del espacio físico, al estudiar la dinámica del tiempo-espacio del Universo aplicando el principio cosmológico débil y el postulado de Weyl.

Introduciendo un nuevo sistema de coordenadas angulares χ, θ, φ la variedad esférica tridimensional se puede representar así, teniendo en cuenta que en este caso la coordenada χ varía entre 0 y 2π ; θ entre 0 y π ; φ entre 0 y 2π :

$x_1 = R \cdot \cos\chi \cdot sen\theta \cdot \cos\varphi$ $\quad$ $x_2 = R \cdot \cos\chi \cdot sen\theta \cdot sen\varphi$ $\quad$ $x_3 = R \cdot \cos\chi \cdot \cos\theta$

$x_4 = R \cdot sen\chi$

Por tanto es una variedad finita, cerrada y en sí misma ilimitada. Es finita porque la distancia entre dos cualesquiera de sus puntos está acotada. Es cerrada e ilimitada porque recorriéndola por adentro no se encuentra una frontera.

En forma análoga se puede definir en E_4 una variedad hiperbólica de 3 dimensiones (**3-hiperbolide de dos hojas**), por la ecuación $x_1^2 + x_2^2 + x_3^2 - x_4^2 = R^2$.

Si por ejemplo tomamos $x_1 = 0$, es decir si nos restringimos a la traza de la variedad sobre un sub-espacio euclidiano tridimensional, obtenemos que dicha traza viene descrita por la ecuación $x_2^2 + x_3^2 - x_4^2 = R^2$ que corresponde a la superficie conocida como hiperboloide de una hoja. Lo mismo ocurre si hacemos $x_2 = 0$ o $x_3 = 0$.

Pero si tomamos $x_4 = 0$ se obtiene la superficie esférica $x_1^2 + x_2^2 + x_3^2 = R^2$.

Introduciendo un nuevo sistema de coordenadas χ, θ, φ la variedad hiperbólica tridimensional se puede representar así, teniendo en cuenta que ahora χ varía entre 0 y $+\infty$, mientras que las coordenadas angulares varían como de costumbre, es decir θ entre 0 y π ; φ entre 0 y 2π :

$x_1 = R \cdot \cosh\chi \cdot sen\theta \cdot \cos\varphi$ $\quad$ $x_2 = R \cdot \cosh\chi \cdot sen\theta \cdot sen\varphi$ $\quad$ $x_3 = R \cdot \cosh\chi \cdot \cos\theta$

$x_4 = R \cdot senh\chi$

Puesto que el coseno hiperbólico y el seno hiperbólico son funciones no acotadas (tienden a $+\infty$ cuando χ tiende a $+\infty$), tenemos una variedad tridimensional que, además de curvada e ilimitada, es un espacio abierto en el siguiente sentido: nos podemos alejar tanto como queramos a partir de un punto, puesto que no hay una cota superior para las distancias entre puntos del espacio.

A su vez la ecuación $x_4^2 - x_1^2 - x_2^2 - x_3^2 = R^2$ define en el espacio E_4 otra variedad hiperbólica de tres dimensiones (**3-hiperboloide de una hoja**), cuya traza para $x_1 = 0$ es $x_4^2 - x_2^2 - x_3^2 = R^2$, que es la ecuación en un espacio euclidiano tridimensional de la superficie conocida como hiperboloide de una hoja.

Finalmente, la ecuación $2 \cdot a \cdot x_4 - x_1^2 - x_2^2 - x_3^2 = 0$ define en E_4 una variedad parabólica tridimensional (**3-parabolide**).

Todas estas variedades tridimensionales se pueden describir en forma intrínseca comenzando por el elemento de línea de E_4 ($ds^2 = (dx_1)^2 + (dx_2)^2 + (dx_3)^2 + (dx_4)^2$), despejando x_4 a partir de la ecuación extrínseca (en E_4) de cada variedad en función de R y de x_1, x_2, x_3 e introduciendo coordenadas esféricas $x_1 = r \cdot \cos\theta \cdot \cos\varphi$, $x_2 = r \cdot \cos\theta \cdot sen\varphi$ y $x_3 = r \cdot sen\theta$.

Se obtiene así, luego de algunos cálculos, una expresión general intrínseca para el elemento de línea de estas variedades tridimensionales curvadas expresado en función de las tres coordenadas r, θ, φ:

$$ds^2 = \left(1 + h^2(r)\right) \cdot dr^2 + r^2 \cdot d\theta^2 + r^2 \cdot sen^2\theta \cdot d\varphi^2$$

Para: la **3-esfera** $h^2(r) = \dfrac{r^2}{R^2 - r^2}$, el **3-hiperbolide de dos hojas** $h^2(r) = \dfrac{r^2}{R^2 + r^2}$, el **3-hiperboloide de una hoja** $h^2(r) = \dfrac{r^2}{r^2 - R^2}$ y el **3-paraboloide** $h^2(r) = \dfrac{r^2}{a^2}$.

En general la inmersión de una variedad tridimensional curva en un espacio euclidiano requiere un espacio de $N = \dfrac{n \cdot (n+1)}{2} = \dfrac{3 \cdot (3+1)}{2} = 6$ dimensiones, o sea E_6.

En consecuencia, las variedades definidas por una ecuación del tipo $F(x_1, x_2, x_3, x_4) = 0$ en E_4 no agotan las variedades tridimensionales posibles.

Para ver esto, comencemos por observar que un elemento de línea de E_4, que adopta la forma $ds^2 = (dx_1)^2 + (dx_2)^2 + (dx_3)^2 + (dx_4)^2$ en coordenadas cartesianas ortogonales, introduciendo coordenadas esféricas en E_3 puede llevarse a la forma siguiente:

$$ds^2 = (dr)^2 + r^2 \cdot (d\theta)^2 + r^2 \cdot sen^2\theta \cdot (d\varphi)^2 + (dx_4)^2$$

Si x_4 se despeja, a partir de $F(x_1, x_2, x_3, x_4) = 0$, en función de x_1, x_2, x_3 y estas últimas tres coordenadas se representan a su vez en coordenadas esféricas, $(dx_4)^2$ aporta un término no negativo de la forma $g^2(r, \theta, \varphi) \cdot (dr)^2$.

Por este motivo el coeficiente que multiplica a $(dr)^2$ en la fórmula para ds^2 es necesariamente mayor o igual a uno.

Un ejemplo de variedad tridimensional curvada que no se puede embeber en E_4 y que, como se verá en el capítulo 12, posee interés cosmológico, es la variedad cuyo elemento de línea viene dado por $ds^2 = \left(\dfrac{R^2}{R^2 + r^2}\right) \cdot dr^2 + r^2 \cdot d\theta^2 + r^2 \cdot sen^2\theta \cdot d\varphi^2$, puesto que el coeficiente $\dfrac{R^2}{R^2 + r^2}$ que multiplica a $(dr)^2$ es inferior a uno. No obstante, esta variedad se puede embeber en un espacio tetra-dimensional no euclidiano[135].

[135] Se la puede embeber en una variedad lineal de cuatro dimensiones con la (pseudo)métrica Lorentziana $ds^2 = (dx_1)^2 + (dx_2)^2 + (dx_3)^2 - (dx_4)^2$, introduciendo la restricción $x_4^2 = R^2 + x_1^2 + x_2^2 + x_3^2$ para caracterizar la variedad curvada como una hiper-superficie en este espacio no euclidiano. Nótese que si se pone $x_4 = 0$ se obtiene un sub-espacio euclidiano E_3 en la variedad lineal no euclidiana. A partir de la restricción y la (pseudo) métrica un cálculo simple introduciendo coordenadas esféricas en el sub-espacio E_3 de la variedad Lorentziana conduce al elemento de línea que se buscaba.

9.4.6 Variedades 3-D con simetría esférica. Representación conforme en E_3.

Finalmente prestemos atención a un tipo de variedad tridimensional (3-D) que posee simetría esférica en torno a uno de sus puntos, designado por O[136] y tomado como origen de un sistema de coordenadas.

El elemento de línea se puede poner así: $ds^2 = g^2(r)\cdot dr^2 + r^2\cdot\left(d\theta^2 + sen^2\theta\cdot d\varphi^2\right)$

$(0 \le r < r_{máx}\,,\ 0 \le \theta \le \pi\,,\ 0 \le \varphi \le 2\pi)$

La función $g(r)$ es regular y no se anula en $0 \le r < r_{máx}$. En este intervalo se la puede tomar positiva. Es posible que $r_{máx}$ sea infinito.

En particular, para la construcción del modelos cosmológicos nos interesan los casos que presentan $g(r) = \dfrac{1}{\sqrt{1-\varepsilon\cdot\dfrac{r^2}{r_0{}^2}}}$ con $\varepsilon = -1$, $\varepsilon = 0$ y $\varepsilon = +1$.

En 9.5.3 estudiaremos algunas propiedades de las curvas geodésicas y en 9.5.4 estudiaremos algunos aspectos de la curvatura de este tipo de variedad no euclidiana.

Ahora, por su importancia en la discusión de los modelos cosmológicos que se realiza en el capítulo 12, veamos la posibilidad de una representación conforme en E_3.

Por definición una correspondencia punto a punto entre dos variedades es conforme cuando los elementos de línea son localmente semejantes: $ds^2 = A^2\cdot ds_1^2$

El factor de semejanza local A puede variar de un punto a otro, siendo en ese caso una función regular de las coordenadas del punto, pero no se anula.

Como el ángulo entre dos curvas que se intersecan en una variedad es igual al ángulo entre sus imágenes en la otra variedad, la representación de una variedad en la otra se denomina conforme.

En el caso particular de las variedades con simetría esférica buscamos una representación conforme con $ds_1^2 = dr_1^2 + r_1{}^2\cdot\left(d\theta^2 + sen^2\theta\cdot d\varphi^2\right)$.

$$g^2(r)\cdot dr^2 + r^2\cdot\left(d\theta^2 + sen^2\theta\cdot d\varphi^2\right) = A^2\cdot\left(dr_1^2 + r_1{}^2\cdot\left(d\theta^2 + sen^2\theta\cdot d\varphi^2\right)\right)$$

Esa representación se puede construir siempre y cuando $g(r)\cdot dr = A\cdot dr_1$ y $r = A\cdot r_1$

Eliminando A se obtiene: $\dfrac{g(r)}{r}\cdot dr = \dfrac{1}{r_1}\cdot dr_1$ Integrando aparece una función que vincula uno a uno r con r_1: $\ln\dfrac{r_1}{r_{1,*}} = \displaystyle\int_{r_*}^{r}\frac{g(r')}{r'}\cdot dr'$

De esta expresión podemos hallar la función $r = r(r_1)$. Entonces, como $r = A\cdot r_1$, el factor de semejanza local se puede expresar así: $A(r_1) = \dfrac{r(r_1)}{r_1}$

Si $g(r) = \dfrac{1}{\sqrt{1-\varepsilon\cdot\dfrac{r^2}{r_0^2}}}$ la integración conduce al factor de semejanza: $A^2(r_1) = \dfrac{1}{1+\varepsilon\cdot\dfrac{r_1^2}{4\cdot r_0^2}}$

---

[136] El caso general de las variedades métricas con simetría esférica puede verse en el libro de T. Levi-Civita "The absolute differential calculus", capítulo 12, Blackie & Son, Glasgow, U.K., 1947.

9.5 Tensores y curvatura.

Puesto que la variedad de cuatro dimensiones necesaria para plantear las ecuaciones del campo gravitatorio no es propiamente una variedad métrica, a diferencia de los 3-hiperboloides, la 3-esfera, el 3-paraboloide y en general las variedades tridimensionales con simetría esférica que consideramos en las secciones 9.4.4 y 9.4.5 precedentes, en esta parte estudiaremos variedades pseudo-métricas, más generales, que abarcan el caso que interesa en la teoría generalizada de la relatividad.

Siguiendo el uso establecido en física, de ahora en adelante haremos referencia a ellas como variedades de Riemann, aunque esta denominación debería restringirse a las variedades provistas de una métrica propiamente dicha.

9.5.1 Tensores[137]

Volviendo ahora a la geometría intrínseca de las variedades de n dimensiones, la propiedad fundamental del elemento de línea es que permanece invariante al cambiar las coordenadas.

Consideremos un cambio de coordenadas: $x^j = f^j\left(x'^1, x'^2, ...x'^n\right)$ para j=1, 2,…, n

Entonces: $dx^i = \sum_{l=1}^{n} \frac{\partial f^i}{\partial x'^l} \cdot dx'^l = \sum_{l'} a^i_{l'} \cdot dx'^l \qquad dx^k = \sum_{m=1}^{n} \frac{\partial f^k}{\partial x'^m} \cdot dx'^m = \sum_{m'} a^i_{m'} \cdot dx'^m$

Las componentes de un vector que se transformen igual que los diferenciales de coordenadas se denominan componentes contra-variantes y se habla de un vector contra-variante o de un tensor de orden uno contra-variante.

Los coeficientes locales $a^{m'}_l = \frac{\partial f^{m'}}{\partial x^l}$ correspondientes al cambio de coordenadas inverso verifican: $\sum_l a^{m'}_l \cdot a^l_{k'} = \delta^{m'}_{k'} \qquad \sum_{l'} a^m_{l'} \cdot a^{l'}_k = \delta^m_k$ siendo $\delta^{m'}_{k'}$ y δ^m_k (deltas de Kronecker) iguales a 1 cuando los índices son iguales e iguales a cero en otro caso.

Consideremos ahora un campo escalar o tensor de orden cero $\phi\left(x^1, x^2, ..., x^n\right)$ cuyo diferencial es $d\phi = \sum_l \frac{\partial \phi}{\partial x^l} \cdot dx^l$

Los escalares, por definición, permanecen invariantes bajo un cambio de coordenadas:

$$d\phi = \sum_l \frac{\partial \phi}{\partial x^l} \cdot dx^l = \sum_{m'} \frac{\partial \phi}{\partial x^{m'}} \cdot dx^{m'} = \sum_l \frac{\partial \phi}{\partial x^l} \cdot \sum_{m'} a^l_{m'} \cdot dx'^m = \sum_{m'} \left(\sum_l \frac{\partial \phi}{\partial x^l} \cdot a^l_{m'} \right) . dx'^m$$

De donde resulta $\frac{\partial \phi}{\partial x^{m'}} = \sum_l \frac{\partial \phi}{\partial x^l} \cdot a^l_{m'}$ Las componentes de un vector que se transforman como las componentes $\frac{\partial \phi}{\partial x^l}$ (es decir, como las componentes del gradiente de un escalar) se denominan componentes co-variantes, y se habla de un vector co-variante o tensor co-variante de orden uno.

---

[137] Un estudio conciso sobre análisis tensorial realizado en el espíritu en que se desarrollan las aplicaciones en la teoría de la gravitación de Einstein puede hallarse, por ejemplo, en L. Santaló, "*Vectores y tensores con sus aplicaciones*", EUDEBA, Buenos Aires, 1969.

De la invariancia del elemento de línea: $\sum_{i,k=1}^{n} g_{i,k} dx^i dx^k = \sum_{l',m'=1}^{n} g_{l'',m'} dx'^l dx'^m$ y de las ecuaciones de transformación de los diferenciales de las coordenadas resulta:

$$g_{l'',m'} = \sum_{i,k=1}^{n} g_{i,k} a_{l'}{}^{i} a_{m'}{}^{k}$$

Un ente que se transforma como las componentes $g_{i,k}\left(x^1, x^2, ..., x^n\right)$del elemento de línea se denomina tensor co-variante de orden dos. En particular el campo tensorial de componentes $g_{i,k}$ que define la métrica o la pseudo-métrica a través del elemento de línea se denomina **tensor métrico de la variedad**. Un tensor muy importante, relacionado con el tensor co-variante de componentes $g_{i,k}$, es el tensor métrico contra-variante cuyas componentes $g^{i,k}$ forman los elementos de la matriz inversa G^{-1} de la matriz G cuyos elementos son los $g_{i,k}$ [138] :

$$\sum_k g^{i,k} \cdot g_{k,j} = \sum g_{j,k} \cdot g^{k,i} = \delta_j^i \qquad (\delta_j^i = 1 \text{ si; i=j ; } \delta_j^i = 0 \text{ si i}\neq\text{j)}$$

Esta propiedad del tensor métrico se utiliza para subir o bajar índices en un tensor.

Así, a partir de las componentes contra-variantes v^j de un vector obtenemos sus componentes co-variantes $v_i = \sum_j g_{i,j} \cdot v^j$, y del tensor en forma mixta (una vez contra-variante y dos veces co-variante) $T_{j,k}{}^{i}$ obtenemos las formas contra-variante $T^{p,q,i} = \sum_{j,k} g^{p,j} \cdot g^{q,k} \cdot T_{j,k}{}^{i}$ y co-variante $T_{p,q,i} = \sum_{j,k} g_{p,j} \cdot g_{q,k} \cdot T_{j,k}{}^{i}$

En cada punto de la variedad, a partir de las leyes de transformación de sus componentes bajo cambios de coordenadas, se pueden definir tensores de orden superior, considerando como un mismo tensor todas las formas que se puedan obtener una de otra subiendo o bajando índices con ayuda del tensor métrico.

En general un tensor de orden m+n que es m-veces contravariante y n-veces covariante se dice que es de orden de variancia (m,n). Por ejemplo, $T_{j,k}{}^{i}$ representan las componentes de un tensor de orden de variancia (1,2), $T^{p,q,i}$ representan las de un tensor de orden de variancia (3,0) y $T_{p,q,i}$ las de uno de orden de variancia (0,3).

El orden en que aparecen los índices covariantes y contravariantes es importante y debe ser mantenido: desde el punto de vista de las operaciones del análisis tensorial, las componentes $T_{j,k}{}^{i}$ no corresponden al mismo objeto matemático que las componentes $T^{i}{}_{j,k}$.

[138] Esa matriz G^{-1} inversa respecto de G existe siempre, puesto que en un espacio de Riemann se supone que $\det\left[g_{i,k}\right] \neq 0$ en todo punto.

Los campos de tensores (digamos $T_{j,k}{}^{i}$) se pueden multiplicar por un escalar (digamos, por a) originando un campo de tensores del mismo orden de variancia y estructura de índices: $a \cdot T_{j,k}{}^{i}$

Los campos tensoriales del mismo orden de variancia y estructura de índices se pueden sumar en cada punto originado un nuevo campo de ese mismo orden: por ejemplo $T_{j,k}{}^{i} + S_{j,k}{}^{i}$ a partir de los tensores de componentes $T_{j,k}{}^{i}$ y $S_{j,k}{}^{i}$.

También se pueden multiplicar las componentes para construir tensores de orden superior: por ejemplo multiplicando las componentes del vector covariante v^{j} de orden (1,0) por las componentes del tensor covariante de orden (0,3) $T_{p,q,i}$, se obtiene el tensor $S^{j}{}_{p,q,i}$ de componentes $v^{j} \cdot T_{p,q,i}$ y de orden de variancia (1,3). Esta operación se denomina producto externo de tensores. No es conmutativo.

Se define un proceso de contracción de índices involucrando un índice covariante y un índice contravariante, que por suma sobre esos índices disminuye en dos el orden del tensor de esta forma: $\sum_{k} S^{k}{}_{p,k,i} = T_{p,i}$

Cuando la contracción de índices se aplica al producto externo de tensores produce un tensor denominado producto interno de los factores: $\sum_{k} v^{k} \cdot T_{p,q,k} = R_{p,q}$

Cuando el número de índices covariantes y contravariantes lo permite, la contracción de índices puede aplicarse a más de un par de índices. En lo que sigue **no recurriremos** a la convención de Einstein, quien propuso eliminar los símbolos de sumatoria asociados a la contracción.

Para determinar si un objeto dado por un conjunto de componentes indexadas es o no un tensor se puede recurrir a un teorema conocido como **ley del cociente**: siempre que el producto interno formal de las componentes del objeto por las componentes de un tensor arbitrario da como resultado las componentes de un tensor, entonces el objeto mismo es un tensor. Así, si dado cualquier tensor $B^{i,j}$ y un mismo objeto $Q^{r}{}_{k,l}$ resulta que $\sum_{s,t} Q^{r}{}_{s,t} \cdot B^{s,t} = C^{r}$ es un vector contravariante, entonces $Q^{r}{}_{k,l}$ es un tensor.

En general las derivadas parciales de las componentes de un campo tensorial no forman un nuevo campo tensorial. Para formarlo se requiere un procedimiento especial de derivación, denominado **derivación covariante**, que se considera en la próxima sección.

Estos campos tensoriales así construidos permiten expresar las leyes físicas en forma independiente de las coordenadas. Como contrapartida la formulación resultante es una formulación local de las leyes. En este sentido el análisis tensorial es la herramienta óptima para formular un marco teórico basado en el principio de contigüidad considerado en el capítulo 3.

9.5.2 Coeficientes de conexión, desplazamiento paralelo y derivadas covariantes

Supongamos que a cada punto $\left(x^1, x^2, ..., x^n\right)$ de una variedad regular V_n le asignamos su espacio tangente $E_n\left(x^1, x^2, ..., x^n\right)$, en el cual seleccionamos una base $\left(\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\right)$ y un vector $\vec{v} = \sum_{j=1}^{n} v^j \cdot \vec{e}_j$.

Consideremos ahora una curva regular γ en la variedad V_n Esa curva viene dada por las funciones $\left(x^1(u), x^2(u), ..., x^n(u)\right)$ del parámetro real u.

Supongamos que a cada punto $\left(x^1(u), x^2(u), ..., x^n(u)\right)$ de la curva γ le hacemos corresponder un vector localizado en la variedad lineal tangente en ese mismo punto:

$$\vec{v}\left(x^1(u), x^2(u), ..., x^n(u)\right) = \sum_{j=1}^{n} v^j\left(x^1(u), x^2(u), ..., x^n(u)\right) \cdot \vec{e}_j\left(x^1(u), x^2(u), ..., x^n(u)\right)$$

Esta expresión puede resumirse poniéndola en términos de funciones del parámetro u:

$\vec{v}[u] = \sum_{j=1}^{n} v^j[u] \cdot \vec{e}_j[u]$ A cada punto de la curva γ se le hace corresponder un vector contenido en la variedad lineal tangente a V_n en ese punto.

Como todas las variedades lineales tangentes son equivalentes, podemos representar $\vec{v}\left(x^1(u), x^2(u), ..., x^n(u)\right)$ como función vectorial en una misma variedad lineal, al igual que los campos vectoriales formados por los vectores de las bases de las variedades lineales tangentes en cada punto de γ:

$$\left\{\vec{e}_1\left(x^1(u), x^2(u), ..., x^n(u)\right), \vec{e}_2\left(x^1(u), x^2(u), ..., x^n(u)\right), ..., \vec{e}_n\left(x^1(u), x^2(u), ..., x^n(u)\right)\right\}$$

Asumimos que los vectores $\vec{e}_j\left(x^1, x^2, ..., x^n\right)$ son funciones regulares de $\left(x^1, x^2, ..., x^n\right)$.

Entonces: $$\frac{\partial}{\partial x^k}\vec{e}_j\left(x^1, x^2, ..., x^n\right) = \sum_{i=1}^{n} \Gamma^i{}_{j,k}\left(x^1, x^2, ..., x^n\right) \cdot \vec{e}_i\left(x^1, x^2, ..., x^n\right)$$

Las n^3 funciones $\Gamma^i{}_{j,k}\left(x^1, x^2, ..., x^n\right)$ se conocen como **coeficientes de conexión**.

Si $d\vec{r} = \sum_{j=1}^{n} dx^j \cdot \vec{e}_j\left(x^1, x^2, ..., x^n\right)$ y $\vec{e}_i \bullet \vec{e}_j = g_{i,j}$, en cada punto de la variedad, el elemento de línea se puede escribir así: $$ds^2 = d\vec{r} \bullet d\vec{r} = \sum_{i,j=1}^{n} \vec{e}_i \bullet \vec{e}_j \cdot dx^i \cdot dx^j$$

Derivando las componentes del tensor métrico: $$\frac{\partial}{\partial x^k} g_{i,j} = \frac{\partial}{\partial x^k}\vec{e}_i \bullet \vec{e}_j + \vec{e}_i \bullet \frac{\partial}{\partial x^k}\vec{e}_j$$

Estas derivadas se pueden expresar en términos de los coeficientes de conexión y de las componentes $\vec{e}_i \bullet \vec{e}_j = g_{i,j}$:

$$\frac{\partial}{\partial x^k} g_{i,j} = \sum_{l=1}^{n} \Gamma^l{}_{i,k} \cdot g_{l,j} + \sum_{l=1}^{n} \Gamma^l{}_{j,k} \cdot g_{l,i}$$

Teniendo en cuenta la definición de las componentes contravariantes $g^{i,j}$ del tensor métrico, vistas en 9.5.1, luego de algunas transformaciones algebraicas se obtiene:

$$\Gamma^{i}{}_{j,k} = \frac{1}{2} \cdot \sum_{l=1}^{n} g^{i,l} \cdot \left(\frac{\partial g_{l,k}}{\partial x^{j}} + \frac{\partial g_{j,l}}{\partial x^{k}} - \frac{\partial g_{j,k}}{\partial x^{l}} \right)$$

Se advierte que los coeficientes de conexión, también conocidos como **símbolos de Christoffel**, quedan determinados por el tensor métrico de la variedad y sus derivadas parciales de primer orden respecto de las coordenadas. Como este tensor es simétrico, los coeficientes de conexión no son independientes[139]. En particular: $\Gamma^{i}{}_{j,k} = \Gamma^{i}{}_{k,j}$

Las componentes $\Gamma_{j,k}{}^{i}$ no forman un tensor.

Dado $\vec{v}[u] = \sum_{j=1}^{n} v^{j}[u] \cdot \vec{e}_{j}[u]$ podemos calcular su derivada respecto del parámetro u.

En primer lugar: $\frac{d}{du}\vec{v}[u] = \sum_{j=1}^{n} (\frac{d}{du} v^{j}[u] \cdot \vec{e}_{j}[u] + v^{j}[u] \cdot \frac{d}{du}\vec{e}_{j}[u])$

Pero sabemos que se verifica:

$$\frac{d}{du} e_{j}[u] = \sum_{k=1}^{n} \frac{dx^{k}(u)}{du} \frac{\partial}{\partial x^{k}} \vec{e}_{j}\left(x^{1}, x^{2}, ..., x^{n}\right) = \sum_{i,k=1}^{n} \Gamma^{i}{}_{j,k}\left(x^{1}, x^{2}, ..., x^{n}\right) \cdot \frac{dx^{k}}{du} \vec{e}_{i}\left(x^{1}, x^{2}, ..., x^{n}\right)$$

En consecuencia obtenemos:

$$\frac{d}{du}\vec{v}[u] = \sum_{j=1}^{n} (\frac{d}{du} v^{j}[u] + \sum_{i,k=1}^{n} \Gamma^{j}{}_{i,k} \cdot v^{i}[u] \cdot \frac{dx^{k}(u)}{du}) \cdot \vec{e}_{j}[u]$$

Las componentes $\frac{\delta}{\delta \cdot u} v^{j}[u] = \frac{d}{du} v^{j}[u] + \sum_{i,k=1}^{n} \Gamma^{j}{}_{i,k} \cdot v^{i}[u] \cdot \frac{dx^{k}(u)}{du}$ $\quad (j = 1,2,...,n)$

definen un vector derivada: $\frac{d}{du}\vec{v}[u] = \sum_{j=1}^{n} (\frac{\delta}{\delta \cdot u} v^{j}[u]) \cdot \vec{e}_{j}[u]$

Los vectores de una variedad lineal se pueden describir mediante una única base de vectores por el origen. En ese caso las componentes del tensor métrico $\vec{e}_{i} \bullet \vec{e}_{j} = g_{i,j}$ son constantes al pasar de un punto de la variedad a otro: la variedad lineal misma es su propia variedad tangente en cada uno de sus puntos.

Si las componentes $g_{i,j}$ son constantes, se anulan idénticamente todos los coeficientes de conexión. Entonces sobre una curva definida en una variedad lineal que se describe mediante una base única, se verifica: $\frac{\delta}{\delta \cdot u} v^{j}[u] = \frac{d}{du} v^{j}[u]$ $\quad (j = 1,2,...,n)$

Si trasladamos un vector $\vec{v}$ paralelo a sí mismo sobre una curva γ contenida en una variedad lineal, en condiciones en las que las componentes del tensor métrico no varían, entonces las componentes del vector respecto de la base tampoco varían:

---

[139] Una exposición detallada de los fundamentos sobre variedades y tensores puede hallarse en el libro de A. Lichnerowicz "Elementos de cálculo tensorial", Aguilar, Madrid, 1968. Para un estudio conciso pero riguroso sobre análisis tensorial, ver por ejemplo, el libro de L. Santaló, "*Vectores y tensores con sus aplicaciones*", EUDEBA, Buenos Aires, 1969. A nivel avanzado se puede consultar el libro, que presenta un enfoque moderno, de D.Lovelock y H.Rund "*Tensors, differential forms and variational principles*" Wiley, New York, 1975, o el clásico de T. Levi-Civita "The absolute differential calculus", Blackie & Son, Glasgow, U.K., 1947.

$$\frac{\delta}{\delta u} v^j[u] = \frac{d}{du} v^j[u] = 0$$

Pero si modificamos la base de vectores al pasar de un punto a otro, el desplazamiento paralelo del vector se produce siempre y cuando $\frac{d}{du}\vec{v}[u] = \sum_{j=1}^{n}(\frac{\delta}{\delta \cdot u} v^j[u]) \cdot \vec{e}_j[u] = 0$ sobre γ, es decir siempre que: $\frac{\delta}{\delta \cdot u} v^j[u] = 0 \qquad (j = 1,2,...,n)$

Es posible generalizar esta caracterización del transporte paralelo de un vector a lo largo de una curva al caso de las variedades riemannianas que hemos estado considerando. Diremos que **transportamos en forma paralela** el vector de componentes $v^j[u]$ sobre la curva γ cuyas coordenadas vienen dadas por $(x^1(u), x^2(u),..., x^n(u))$ siempre y cuando **a lo largo de esa curva** se verifiquen las ecuaciones: $\frac{\delta}{\delta \cdot u} v^j[u] = 0 \qquad (j = 1,2,...,n)$

En forma equivalente: $\frac{d}{du} v^j[u] = -\sum_{i,k=1}^{n} \Gamma^j{}_{i,k} \cdot v^i[u] \cdot \frac{dx^k(u)}{du} \qquad (j = 1,2,...,n)$

Teniendo en cuenta que $\frac{d}{du} v^j[u] = \sum_{k=1}^{n} \frac{\partial}{\partial x^k} v^j(x^1(u), x^2(u),..., x^n(u)) \cdot \frac{dx^k(u)}{du}$, la ecuación del transporte paralelo se verifica si a lo largo de la curva γ:

$$\frac{\partial}{x^k} v^j(x^1(u), x^2(u),..., x^n(u)) + \sum_{i=1}^{n} \Gamma^j{}_{i,k}(x^1(u), x^2(u),..., x^n(u)) \cdot v^i(x^1(u), x^2(u),..., x^n(u)) = 0$$

$$(j, k = 1,2,...,n)$$

En esta ecuación ha aparecido el siguiente tensor mixto de segundo orden, denominado **derivada covariante** del vector de coordenadas v^j:

$$D_k v^j(x^1, x^2,..., x^n) = \frac{\partial}{x^k} v^j(x^1, x^2,..., x^n) + \sum_{i=1}^{n} \Gamma^j{}_{i,k}(x^1, x^2,..., x^n) \cdot v^i(x^1, x^2,..., x^n)$$

Entonces una condición que asegura el cumplimiento de la ecuación del transporte paralelo del vector a lo largo de la curva γ se puede resumir estableciendo que sobre esa curva se anula la derivada covariante $D_k v^j(x^1(u), x^2(u),..., x^n(u)) = 0$

La derivación covariante se introduce debido a que las derivadas parciales de las componentes de un campo tensorial no constituyen las componentes de otro tensor, con excepción del caso en el que las variedades son euclidianas o pseudo-euclidianas y se describen mediante sistemas de coordenadas cartesianas ortogonales.

Generaliza la derivación parcial y permite construir nuevos campos de tensores.

Se puede definir para tensores de cualquier orden de covariancia o contravariancia, empleando los símbolos de Christoffel y las componentes del tensor métrico correspondiente a la variedad sobre la que se trabaja.

Para un campo tensorial dos veces contravariante la derivada covariante resulta ser un tensor de orden tres, una vez covariante y dos veces contravariante:

$$D_k T^{i,j} = \frac{\partial}{x^k} T^{i,j} + \sum_{l=1}^{n} \Gamma^j{}_{k,l} \cdot T^{l,j} + \sum_{l=1}^{n} \Gamma^j{}_{k,l} \cdot T^{i,l}$$

De esta fórmula, junto con la relación $\sum_{m=1}^{n} g^{i,m} \cdot g_{m,j} = \delta_j^i$ y la definición de los coeficientes de Christoffel en términos de las componentes del tensor métrico se deduce que la derivada covariante $D_k g^{i,j}$ del tensor métrico es nula en todos los puntos de la variedad: $D_k g^{i,j}\left(x^1, x^2, ..., x^n\right) = 0$

De esta propiedad del tensor métrico y la definición de transporte paralelo de un vector (de componentes v^j) sobre una curva, se desprende que $\sum_{i,j=1}^{n} g_{i,j} \cdot v^i \cdot v^j$ no varía con el transporte paralelo. Cuando la variedad posee una métrica propiamente dicha, $\sum_{i,j=1}^{n} g_{i,j} \cdot v^i \cdot v^j$ puede interpretarse como el cuadrado de la longitud del vector, la cual, en consecuencia, no varía durante el transporte paralelo.

Efectuando una contracción de índices a partir de la derivada covariante $D_k T^{i,j}$ del tensor escrito en forma contravariante $T^{i,j}$, obtenemos el tensor $\sum_{l=0}^{3} D_l T^{l,j}$ denominado **divergencia covariante**.

Cuando los coeficientes de conexión $\Gamma_{j,k}{}^{i}$ son nulos, la derivada covariante se reduce a $\frac{\partial}{\partial x^k} T^{i,j}$ y la divergencia covariante se reduce a $\sum_{l=0}^{3} \frac{\partial}{\partial x^l} T^{l,j}$.

Como se verá en la sección 9.5.4, cuando los coeficientes de conexión se anulan en todos los puntos de la variedad, ésta posee curvatura nula: se dice que es una variedad plana. El tiempo-espacio de Minkowski de la relatividad restringida es un ejemplo de variedad plana, en el cual $\frac{\partial}{\partial x^k} T^{i,j}$ es un tensor de orden de variancia (2,1) y $\sum_{l=0}^{3} \frac{\partial}{\partial x^l} T^{l,j}$ es un tensor (1,0), es decir, un vector contravariante.

9.5.3 Curvas geodésicas

En una variedad con una métrica propiamente dicha, el rol que tienen las rectas del espacio euclidiano lo tienen las líneas geodésicas, que unen dos puntos lo suficientemente próximos mediante el arco de curva de longitud mínima.

Teniendo en cuenta que $ds^2 = \sum_{i,k=1}^{n} g_{i,k} dx^i dx^k$, la longitud del arco $\gamma_{A,B}$ de la curva γ que une los puntos A y B de parámetros u_A y u_B viene dada por:

$$s\left(\gamma_{A,B}\right) = \int_{u_A}^{u_B} \sqrt{\sum_{i,k=1}^{n} g_{i,k} \frac{dx^i}{du} \frac{dx^k}{du}} \cdot du$$

Aplicando las herramientas del cálculo de variaciones[140], se obtiene el siguiente sistema de ecuaciones diferenciales ordinarias, en general no lineales, cuyas soluciones permiten

[140] Para una brevísima introducción al cálculo de variaciones, incluyendo las ecuaciones de Euler (de las cuales las ecuaciones de las geodésicas constituyen un caso particular), ver la sección 11.4.1. Para un

la determinación de las geodésicas que pasan por cada punto de la variedad y en cualquier dirección:

$$\frac{d^2x^i(s)}{du^2}+\sum_{j,k}\Gamma_{j,k}{}^i\left(x^1(u),x^2(u),...,x^n(u)\right)\cdot\frac{dx^j(u)}{du}\cdot\frac{dx^k(u)}{du}=0 \qquad i=1,2,...,n$$

Como condiciones iniciales se requiere especificar un punto de la variedad y una dirección, dada por las componentes de un vector que entonces va resultar tangente a la línea geodésica en el punto inicial.

Consideremos, a vía de ejemplo, una variedad tridimensional con simetría esférica, del tipo introducido al final de la sección 9.4.

El elemento de línea viene dado por: $ds^2=g^2(r)\cdot dr^2+r^2\cdot d\theta^2+r^2\cdot sen^2\theta\cdot d\varphi^2$

Supondremos que la función regular $g(r)$ es siempre positiva, con $g(0)=1$, $\frac{d}{dr}g(0)=0$ y $\frac{d^2}{dr^2}g(0)\neq 0$. La longitud de un arco de curva $\gamma_{A,B}$ de extremos A y B, con parámetro u, y contenido en esta variedad viene dada por:

$$s(\gamma_{A,B})=\int_{u_A}^{u_B}\sqrt{g^2(r)\cdot\left(\frac{dr}{du}\right)^2+r^2\cdot\left(\frac{d\theta}{du}\right)^2+r^2\cdot sen^2\theta\cdot\left(\frac{d\varphi}{du}\right)^2}\cdot du$$

Las ecuaciones de las líneas geodésicas para una variedad como ésta se pueden obtener directamente aplicando el método que se explica en el capítulo 11 (sección 11.4.1).

Introducimos la siguiente función, donde $\dot{r}=\frac{dr}{du}$ $\dot{\theta}=\frac{d\theta}{du}$ $\dot{\varphi}=\frac{d\varphi}{du}$:

$$F(r,\theta,\varphi,\dot{r},\dot{\theta},\dot{\varphi})=\sqrt{g^2(r)\cdot\dot{r}^2+r^2\cdot\dot{\theta}^2+r^2\cdot sen^2\theta\cdot\dot{\varphi}^2}$$

Entonces las ecuaciones de Euler para las geodésicas se pueden escribir así:

$$\frac{d}{du}\left(\frac{g^2(r)}{F}\cdot\dot{r}\right)=\frac{g(r)\cdot g'(r)\cdot\dot{r}^2+r\cdot\left(\dot{\theta}^2+sen^2\theta\cdot\dot{\varphi}^2\right)}{F}$$

$$\frac{d}{du}\left(\frac{r^2}{F}\cdot\dot{\theta}\right)=\frac{r\cdot sen\theta\cdot\cos\theta}{F}\cdot\dot{\varphi}^2 \qquad \frac{d}{du}\left(\frac{r^2\cdot sen^2\theta}{F}\cdot\dot{\varphi}\right)=0$$

Del estudio de estas ecuaciones se desprende que las curvas caracterizadas por $\theta=cte$, $\varphi=cte$ y r variable entre 0 y ∞ son todas ellas geodésicas de la variedad.

Si ponemos $r=cte$ y elegimos dos puntos A y B sobre la superficie así determinada, la longitud de un arco de curva que une A con B sin abandonar dicha superficie viene dado por:

$$s(\gamma_{A,B})=r\cdot\int_{u_A}^{u_B}\sqrt{\left(\frac{d\theta}{du}\right)^2+sen^2\theta\cdot\left(\frac{d\varphi}{du}\right)^2}\cdot du$$

Entonces el problema de hallar las curvas geodésicas sobre la superficie en cuestión se reduce al de hallar las geodésicas sobre una esfera, que son arcos de círculo máximo.

No obstante, estas superficies $r=cte$ (pese a que se denominan esferas geodésicas) **no están formadas por curvas geodésicas de la variedad tridimensional: no son superficies geodésicas**.

---

desarrollo más detallado, desde una perspectiva física, puede consultarse el libro de J.L. Basdevant "*Variational principles in Physics*", Springer Verlag, New York, 2007.

Para verlo alcanza con poner $r = cte$ en las ecuaciones de las geodésicas. La primera de ellas se reduce a $\dot{\theta}^2 + sen^2\theta \cdot \dot{\varphi}^2 = 0$ Si $sen\theta$ no se anula, la única solución posible es $\theta = cte$ y $\varphi = cte$. Junto con $r = cte$ determinan un punto en la esfera geodésica y no una curva contenida en dicha esfera. Si $sen\theta = 0$ entonces $\theta = 0$ o $\theta = \pi$. En ese caso tenemos solo dos puntos, los polos de la esfera.
En general, para determinar las geodésicas hay que resolver el sistema de ecuaciones diferenciales no lineales que las caracteriza como curvas de longitud de arco mínima.
De la tercera ecuación para las geodésicas se desprende que $\dfrac{r^2 \cdot sen^2\theta}{F} \cdot \dot{\varphi} = cte$
Si ponemos $\varphi = cte$, las primeras dos ecuaciones se reducen a un sistema de ecuaciones diferenciales no lineales en términos de r y θ solamente. Las soluciones son geodésicas contenidas en la superficie $\varphi = cte$, la cual es entonces una **superficie geodésica**. Utilizaremos este resultado en 9.5.4 para calcular las curvaturas de las variedades simétricas 3-D.
Ahora especificaremos los símbolos de Christoffel en un punto cualquiera de una de estas variedades, desarrollando las derivadas indicadas en las ecuaciones de Euler.
Si tomamos la longitud de arco como parámetro y tenemos en cuenta que en este caso $F = 1$, las ecuaciones de las geodésicas se pueden poner bajo la forma:

$$\frac{d^2r}{ds^2} + \Gamma^r{}_{\theta\theta} \cdot \left(\frac{d\theta}{ds}\right)^2 + \Gamma^r{}_{\varphi\varphi} \cdot \left(\frac{d\varphi}{ds}\right)^2 = 0 \qquad \Gamma^r{}_{\theta\theta} = -\frac{r}{g^2(r)} \qquad \Gamma^r{}_{\varphi\varphi} = -\frac{r \cdot sen^2\theta}{g^2(r)}$$

$$\frac{d^2\theta}{ds^2} + 2 \cdot \Gamma^\theta{}_{\theta,r} \cdot \left(\frac{d\theta}{ds}\right) \cdot \left(\frac{dr}{ds}\right) + \Gamma^\theta{}_{\varphi\varphi} \cdot \left(\frac{d\varphi}{ds}\right)^2 = 0 \qquad \Gamma^\theta{}_{\theta,r} = -\frac{1}{r} \qquad \Gamma^\theta{}_{\varphi\varphi} = -sen\theta \cdot \cos\theta$$

$$\frac{d^2\varphi}{ds^2} + 2 \cdot \Gamma^\varphi{}_{r\varphi} \cdot \left(\frac{dr}{ds}\right) \cdot \left(\frac{d\varphi}{ds}\right) + 2 \cdot \Gamma^\varphi{}_{\varphi\theta} \cdot \left(\frac{d\varphi}{ds}\right) \cdot \left(\frac{d\theta}{ds}\right) = 0 \qquad \Gamma^\varphi{}_{r\varphi} = \frac{1}{r} \qquad \Gamma^\varphi{}_{\varphi\theta} = ctg\theta$$

Los demás símbolos de Christoffel no relacionados con los anteriores por relaciones de simetría, son nulos.

Cuando no existe una métrica propiamente dicha, como ocurre con las variedades que aparecen en la teoría generalizada de la relatividad, es necesario definir las líneas geodésicas de otra manera, sin recurrir a la longitud de arco.
Esta definición se puede hacer recurriendo al desplazamiento paralelo.
Definiremos una **curva geodésica** como aquella cuyos vectores tangentes $v^j[u] = \dfrac{dx^j(u)}{du}$ resultan de un proceso de transporte paralelo de un vector, que define una dirección inicial en un punto inicial, a lo largo de la curva.
Si $v^j[u]$ son las componentes del vector que sufre el transporte paralelo, entonces sabemos que $\dfrac{d}{du} v^j[u] = -\sum_{i,k=1}^{n} \Gamma^j{}_{i,k} \cdot v^i[u] \cdot \dfrac{dx^k(u)}{du}$
Sustituyendo en esta ecuación $v^j[u] = \dfrac{dx^j(u)}{du}$ y reordenando se obtiene en este caso general la misma ecuación de las geodésicas obtenida en el caso particular en el que la variedad posee una métrica propiamente dicha.

Puesto que se basa en transportar en forma paralela, sobre la curva, el vector que caracteriza la dirección en cada punto de esa misma curva, podemos resumir entonces la definición de una curva geodésica diciendo que **una geodésica es una línea auto-paralela**.

9.5.4 Curvatura de una variedad de n dimensiones

Consideremos nuevamente las variedades tridimensionales cuyos elementos de línea vienen dados por expresiones del tipo $ds^2 = g^2(r)\cdot dr^2 + r^2\cdot d\theta^2 + r^2\cdot sen^2\theta\cdot d\varphi^2$ donde la función regular $g(r)$ es siempre positiva con $g(0)=1$, $\frac{d}{dr}g(0)=0$ y $\frac{d^2}{dr^2}g(0)\neq 0$. Introduciendo el radio geodésico $\rho=\int_0^r g(\xi)\cdot d\xi$, el elemento de línea de esas variedades se puede expresar así: $ds^2 = d\rho^2 + r^2(\rho)\cdot\left(d\theta^2 + sen^2\theta\cdot d\varphi^2\right)$ donde $r(\rho)$ es la función inversa de $\rho=\int_0^r g(\xi)\cdot d\xi$.

Supongamos que fijamos los valores del radio geodésico y del ángulo θ. En ese caso, si ponemos $\theta=\frac{\pi}{2}$ y si variamos el ángulo φ entre 0 y 2π, obtenemos una circunferencia geodésica de longitud $l_c(\rho)=\int_0^{2\pi} ds(\varphi)=2\pi\cdot r(\rho)$

Desarrollando en potencias del radio geodésico, siendo $o(\rho^3)$ un resto que tiende a cero con ρ por lo menos como ρ^4:

$$r(\rho)=r(0)+\frac{dr}{d\rho}(0)\cdot\rho+\frac{1}{2}\cdot\frac{d^2r}{d\rho^2}(0)\cdot\rho^2+\frac{1}{6}\cdot\frac{d^3r}{d\rho^3}(0)\cdot\rho^3+o(\rho^3)$$

Calculando los coeficientes del desarrollo de $r(\rho)$ en potencias del radio geodésico se obtiene:

$$r(0)=0 \qquad \frac{dr}{d\rho}(0)=\frac{1}{g(r(0))}=1 \qquad \frac{d^2r}{d\rho^2}(0)=-\frac{1}{g^3(r(0))}\cdot\frac{dg(r(0))}{dr}=0$$

$$\frac{d^3r}{d\rho^3}(0)=-\frac{1}{g^7(r(0))}\cdot\left(g^3(r(0))\cdot\frac{d^2g(r(0))}{dr^2}-3\cdot g^2(r(0))\cdot\left(\frac{dg(r(0))}{dr}\right)^2\right)=-\frac{d^2g(r(0))}{dr^2}$$

Entonces resulta la fórmula que relaciona la longitud de una circunferencia con el radio geodésico[141] y con la curvatura de Gauss introducida en 9.1.2, con la salvedad que allí designamos con r el radio geodésico que aquí hemos representado por ρ:

$$l_c(\rho)=2\pi\cdot r(\rho)=2\pi\cdot\rho-\frac{\pi}{6}\cdot K\cdot\rho^3+\varepsilon_c$$

Como en esta ecuación $K=\frac{d^2g(0)}{dr^2}$, hemos hallado la curvatura de Gauss de la variedad en el punto $r=0$ a partir del coeficiente $g(r)$ de la métrica. Para todas las circunferencias geodésicas con centro en $r=0$ se obtiene la misma fórmula, de modo

[141] Conocida como circunferencia geodésica, aunque en general ella misma no es una geodésica.

que en ese punto la curvatura de la variedad tridimensional se describe mediante un único escalar.
Este método se puede extender para determinar la curvatura en cualquier otro punto.
No obstante, las dificultades del cálculo no son mucho menores que las implicadas en un abordaje general del problema de la curvatura, de modo que conviene analizar este último, comenzando por resumir el punto de vista de Riemann.

Riemann investigó la generalización de la curvatura de Gauss a una variedad n-dimensional y halló que se podía hacer de la siguiente forma.
Se elige el punto P de la variedad en el cual se desea definir la curvatura intrínseca. Luego se considera la totalidad de las líneas geodésicas que pasan por P.
Se construyen las variedades bidimensionales regulares con las geodésicas que pasan por P. Como estas variedades equivalen a las superficies de la teoría de Gauss, se les puede hacer corresponder a cada en el punto P su curvatura de Gauss.
Apliquemos esta idea en un punto P de la variedad 3-D simétrica, distinto del origen O.
La geodésica radial que une O con P es un eje de simetría, de modo que todas las superficies geodésicas que contengan a dicha radial deben poseer la misma curvatura de Gauss en P. Como vimos en 9.5.3, la superficie $\varphi = cte$ es una superficie geodésica. Como P se caracteriza por tres coordenadas r_P, θ_P, φ_P, eligiendo $\varphi = \varphi_P$ la correspondiente superficie geodésica contiene a la geodésica radial por P.
El elemento de línea sobre esta superficie es $ds^2 = E(r,\theta)\cdot dr^2 + G(r,\theta)\cdot d\theta^2$ con $E(r,\theta) = g^2(r)$ $G(r,\theta) = r^2$. Entonces la curvatura de Gauss en un punto de esta superficie se puede hallar aplicando la fórmula vista en 9.4.1:

$$K(r,\theta) = -\frac{1}{\sqrt{E\cdot G}}\left[\frac{\partial}{\partial r}\left(\frac{1}{\sqrt{E}}\cdot\frac{\partial\sqrt{G}}{\partial r}\right) + \frac{\partial}{\partial\theta}\left(\frac{1}{\sqrt{G}}\cdot\frac{\partial\sqrt{E}}{\partial\theta}\right)\right] = -\frac{1}{2\cdot r}\cdot\frac{d}{dr}\left(\frac{1}{g^2(r)}\right)$$

Basta con poner $r = r_P$ para obtener en el punto P las curvaturas de todas las superficies geodésicas que pasa por P y contienen a la geodésica radial que une O con P.
Ahora lo único que falta es determinar la curvatura de una superficie geodésica normal a la geodésica radial en el punto P. La esfera $r = r_P$ es ortogonal en P a dicha geodésica, pero como vimos en 9.5.3, no es una superficie geodésica. Por este motivo la nueva curvatura K_1 que estamos buscando no coincide con la curvatura de Gauss $\frac{1}{r_P^2}$ de la esfera en el punto P.
En cualquier punto de la variedad distinto del origen K_1 depende solamente del valor de la coordenada radial del punto y verifica[142]: $K_1 = \frac{1}{r^2}\cdot\left(1 - \frac{1}{g^2(r)}\right)$

Para poder relacionar entre sí las curvaturas de las distintas superficies geodésicas (variedades bidimensionales formadas por geodésicas) que pasan por un mismo punto

[142] Ver, por ejemplo, el libro de T. Levi-Civita "*The absolute differential calculus*", pp. 414-419, Blackie & Son, Glasgow, U.K., 1947.

de V_n, Riemann introdujo un tensor de orden cuatro $R^l{}_{i,k,j}$ formado a partir de las componentes del tensor métrico y sus derivadas parciales primeras y segundas respecto de las coordenadas[143] (las derivadas segundas aparecen porque es necesario derivar los símbolos de Christoffel). Posteriormente estudiaremos sus componentes. Antes de hacerlo, veamos algunos resultados de interés.

Consideremos las componentes covariantes R_{ljhk} del tensor de curvatura.

Para cada punto P de V_n y para cada par de direcciones por P, dadas por las componentes contravariantes a^j y b^j de dos vectores en la variedad lineal tangente $T_n(P)$, se puede definir una curvatura escalar[144]:

$$K\left(x^1, x^2, ..., x^n; \vec{a}, \vec{b}\right) = \frac{\sum_{l,j,h,k=1}^{n} R_{ljhk} \cdot a^l \cdot a^h \cdot b^j \cdot b^k}{\sum_{l,j,h,k=1}^{n} \left(g_{lh} \cdot g_{jk} - g_{lk} \cdot g_{jh}\right) \cdot a^l \cdot a^h \cdot b^j \cdot b^k}$$

Esta curvatura no cambia si los vectores $\vec{a}$ y $\vec{b}$ se sustituyen por otros dos, linealmente independientes, obtenidos por combinación lineal de $\vec{a}$ y $\vec{b}$. Como consecuencia puede asignarse en el punto P a la superficie geodésica $V_{g,2}$ tangente al 2-plano $T_2(P)$ generado por esos vectores. Además, se puede demostrar que coincide con la curvatura Gaussiana de la superficie $V_{g,2}$ en el punto P, calculada con los métodos de la geometría diferencial clásica.

Las componentes de $R^l{}_{i,k,j}$ se anulan en un punto siempre y cuando la variedad sea localmente plana, en el sentido de que todas las superficies geodésicas que pasan por el punto posean curvatura de Gauss nula. Esto ocurre si y solo si existe un sistema de coordenadas cartesianas $\left(y^1, y^2, ..., y^n\right)$ en el cual el elemento de línea se reduce a la forma $ds^2 = \sum_{l=1}^{n} \varepsilon_l \cdot dy_l^2$ donde cada ε_l vale ± 1.

A partir del tensor de curvatura de Riemann- Christoffel, sumando las componentes $R^l{}_{i,k,l}$, Ricci construyó un tensor de curvatura simétrico de orden dos, denominado tensor de Ricci: $R_{i,k} = \sum_{l=1}^{n} R^l{}_{i,k,l}$

La suma de las componentes que poseen un índice contra-variante y un índice co-variante iguales se denomina **contracción del tensor**.

A partir del tensor de Ricci, mediante una doble contracción se obtiene una curvatura escalar $R = \sum_{i,k=1}^{n} g^{i,k} \cdot R_{i,k}$ Puede considerarse como una especie de generalización a una

[143] Conocido como tensor de curvatura de Riemann- Christoffel.

[144] Ver, por ejemplo, el libro de D.Lovelock y H.Rund "*Tensors, differential forms and variational principles*", pp. 264-266, Wiley, New York, 1975

variedad V_n de la curvatura escalar de Gauss. Tanto el tensor de Ricci $R_{i,k}$ como R aparecen en las ecuaciones de la gravitación de Einstein.

Es posible demostrar que un espacio de Riemann es homogéneo siempre y cuando las componentes del tensor de curvatura sean constantes y se reduzcan al mismo valor escalar en todos los puntos de la variedad: $R^{l}{}_{i,k,l} = R \cdot \delta^{l}{}_{i,k,l}$

Siendo $\delta^{l}{}_{i,k,l} = 1$ siempre que $l = i = j = k$; y $\delta^{l}{}_{i,k,l} = 0$ en otro caso.

Es decir, la curvatura, en lugar de venir caracterizada por un tensor de curvatura que cambia de un punto a otro, viene caracterizada por un número, el mismo para todos los puntos.

Si $R = 0$ el espacio homogéneo posee una geometría intrínseca euclidiana o pseudo-euclidiana.

Si R es positivo posee la misma geometría intrínseca que una n-esfera inmersa en un espacio euclidiano o pseudo-euclidiano.

Si R es negativo la geometría intrínseca es del tipo de Lobachevski.

El tensor de curvatura de Riemann-Christoffel aparece en forma natural cuando se efectúa el transporte paralelo de un vector sobre una curva cerrada.

Si la variedad posee una curvatura intrínseca no nula, el vector que se obtiene por transporte paralelo al regresar al punto inicial es en general distinto al vector con el cual comienza el proceso. La figura 9.7 sugiere lo que ocurre al transportar paralelamente un vector: la parte situada a la izquierda corresponde al desplazamiento sobre una curva cerrada en la superficie de un cilindro (que posee curvatura intrínseca, de Gauss, nula), mientras que la parte situada a la derecha corresponde al desplazamiento sobre una curva cerrada en la superficie de una esfera (que posee curvatura intrínseca positiva).

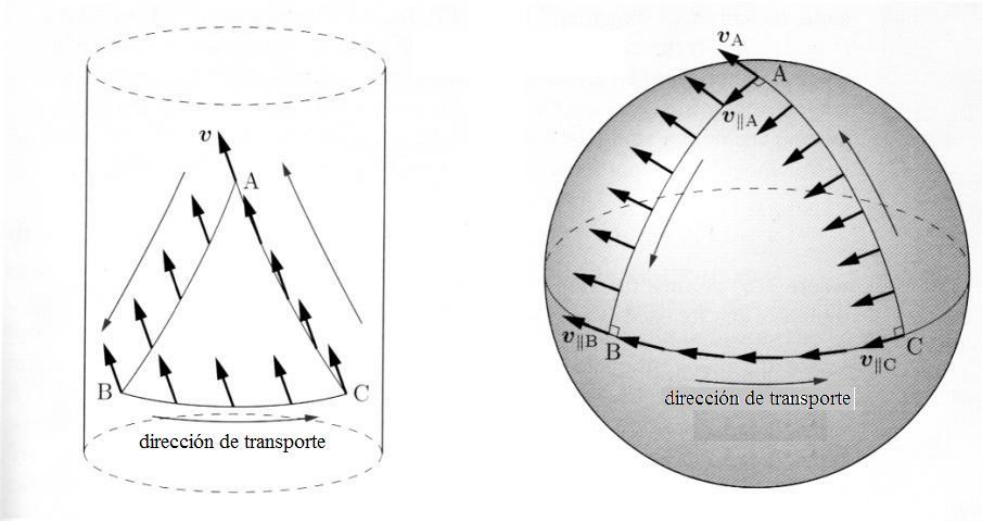


Figura 9.7

En el caso de la variedad cilíndrica el vector $\vec{v}_{\Uparrow A}$ que se obtiene en el punto A al finalizar el transporte coincide con el vector inicial $\vec{v}_A$ en dicho punto. Pero para la variedad esférica esos vectores ya no son iguales.

Para obtener una expresión matemática de lo que acontece se puede analizar el transporte paralelo de un vector inicial $\vec{v}_P$ sobre un circuito elemental $P \to Q \to R \to S \to P$ como el que muestra la Figura 9.8.

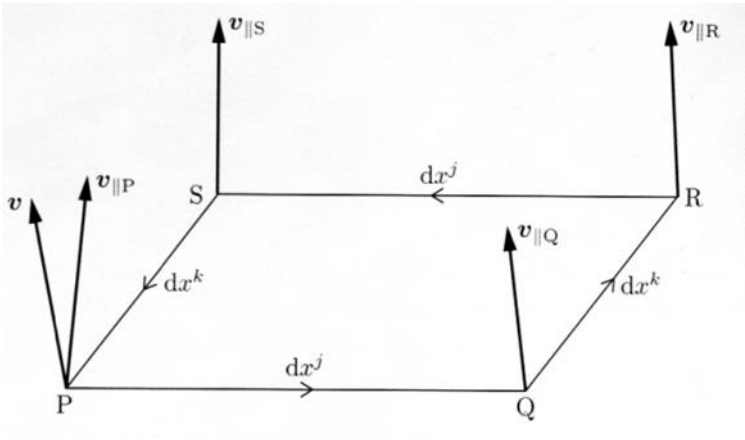


Figura 9.8

De los resultados deducidos en 9.5.2 se desprende la ecuación $dv^l = -\sum_{i,k=1}^{n} \Gamma^l{}_{i,k} \cdot v^i \cdot dx^k$ que indica cómo varían las componentes de los vectores en un desplazamiento paralelo infinitesimal.

A partir de esta ecuación se puede deducir la siguiente relación para la diferencia en un punto P entre cada una de las componentes del campo vectorial construido a partir de un vector inicial en P luego de un proceso de transporte paralelo siguiendo un circuito cerrado O y las componentes correspondientes del vector inicial:

$$v^l{}_{\Uparrow P} - v^l{}_P = \oint_O dv^l = -\oint_O \sum_{i,k=1}^{n} \Gamma^l{}_{i,k} \cdot v^i \cdot dx^k$$

La integral curvilínea $\oint_O dv^l$ se toma sobre el circuito y se puede igualar a una integral ordinaria respecto del parámetro u sobre un intervalo $[u_0, u_1]$ cuyos extremos corresponden al punto P en el que comienza y termina el proceso de integración.

En el límite, cuando un circuito construido como sugiere la Figura 9.8 se contrae hacia el punto P, a partir de la integral de línea sobre la curva cerrada se desprende que:

$$v^l{}_{\Uparrow P} - v^l{}_P \approx \left(\sum_{i=1}^{n} R^l{}_{i,j,k} \cdot v^i \right)_P \cdot dx^j \cdot dx^k$$

En esta expresión las componentes $R^l{}_{i,k,j}$ están referidas al punto P, mientras que las componentes v^i son las del vector inicial $\vec{v}_P$. Durante el cálculo se obtiene:

$$R^l{}_{i,k,j} = \frac{\partial \Gamma^l{}_{i,k}}{\partial x^j} - \frac{\partial \Gamma^l{}_{i,j}}{\partial x^k} + \sum_{m=1}^{n} \Gamma^l{}_{m,j} \cdot \Gamma^m{}_{i,k} - \sum_{m=1}^{n} \Gamma^l{}_{m,k} \cdot \Gamma^m{}_{i,j}$$

Este es el tensor de curvatura de Riemann-Christoffel expresado en términos de los coeficientes de conexión de la variedad.

Como en 9.5.2 se mostró que para las variedades de Riemann los coeficientes de conexión verifican $\Gamma^{i}{}_{j,k} = \frac{1}{2} \cdot \sum_{l=1}^{n} g^{i,l} \cdot \left(\frac{\partial g_{l,k}}{\partial x^{j}} + \frac{\partial g_{j,l}}{\partial x^{k}} - \frac{\partial g_{j,k}}{\partial x^{l}} \right)$, resulta que las $R^{l}{}_{i,k,j}$ son funciones lineales de las derivadas parciales de orden dos y no lineales de las derivadas parciales de orden uno y de las componentes del tensor métrico.
Debido a las simetrías que presentan los coeficientes de conexión y sus derivadas respecto de las coordenadas cuando se expresan en función del tensor métrico, el tensor de Ricci $R_{i,k} = \sum_{l=1}^{n} R^{l}{}_{i,k,l}$ resulta ser simétrico con $\frac{1}{2} \cdot n \cdot (n+1)$ componentes independientes y el tensor de curvatura $R^{l}{}_{i,k,j}$ posee solamente $\frac{1}{12} \cdot n^{2} \cdot (n^{2}-1)$ componentes independientes. En componentes covariantes:

$$R_{ijkm} = -R_{jikm} = -R_{ijmk} \qquad R_{ijkm} + R_{ikmj} + R_{imjk} = 0 \qquad i, j, k, m = 1,2,...,n$$

Si $n = 2$ el tensor de curvatura posee solo una componente independiente, proporcional a la curvatura de Gauss. En forma covariante: $R_{1212} = \sqrt{g_{11} \cdot g_{22} - g_{12}{}^{2}} \cdot K$
Cuando $n = 4$, que es el caso que interesa en la teoría de la gravitación, el tensor de curvatura posee 20 componentes independientes y el tensor de Ricci posee solamente 10.

Un punto P de coordenadas $(x^{1}, x^{2},..., x^{n})$ perteneciente a una variedad riemanniana de dimensión n>2 se dice que es isótropo cuando las curvaturas de Gauss de todas las variedades bidimensionales formadas por geodésicas que pasan por P son iguales. Representemos por $K(x^{1}, x^{2},..., x^{n})$ esa curvatura común. Es posible demostrar que cuando esto ocurre, las componentes covariantes del tensor de curvatura verifican:

$$R_{ljhk} = K(x^{1}, x^{2},..., x^{n}) \cdot (g_{lh} \cdot g_{jk} - g_{lk} \cdot g_{jh})$$

Si todos los puntos del espacio son isótropos, un teorema debido a Schur afirma que entonces el espacio es homogéneo. La curvatura K **es la misma** en todos los puntos de la variedad. Un único escalar caracteriza la curvatura en todos los puntos de la variedad.
En ese caso mediante un cambio de coordenadas adecuado el elemento de línea se puede llevar a la forma canónica[145]: $ds^{2} = \dfrac{\sum_{i=1}^{n} dx_{i}^{2}}{1 + \frac{1}{4} \cdot K \cdot \sum_{i=1}^{n} x_{i}^{2}}$

Podemos dar un ejemplo concreto si consideramos nuevamente las variedades simétricas 3-D e investigamos en qué condiciones todos sus puntos son isótropos.
Para ello debe verificarse la igualdad de las curvaturas para cada r, es decir:

$$K_{1} = \frac{1}{r^{2}} \cdot \left(1 - \frac{1}{g^{2}(r)} \right) = -\frac{1}{2 \cdot r} \cdot \frac{d}{dr} \left(\frac{1}{g^{2}(r)} \right) = K$$

Poniendo $h(r) = \frac{1}{g^{2}(r)}$ y resolviendo la ecuación diferencial resultante, se obtiene $h(r) = 1 - C \cdot r^{2}$, donde la constante de integración C se puede escribir, sin pérdida de

[145] T. Levi-Civita "The absolute differential calculus", capítulo 8, Blackie & Son, Glasgow, U.K., 1947.

generalidad así: $C = \frac{\varepsilon}{r_0^2}$ donde ε puede tomar los valores $-1,0,1$ y r_0 es una constante positiva arbitraria. Entonces la componente $g_{rr} = g(r)$ del tensor métrico de la variedad adopta la forma $g(r) = \left(1 - \varepsilon \cdot \frac{r^2}{r_0^2}\right)^{-\frac{1}{2}}$

Sustituyendo esta función en la expresión para la curvatura de Gauss en un punto resulta que la curvatura es la misma en todos los puntos (es decir, la variedad simétrica 3-D es además de isótropa, homogénea):

$$K = K_1 = \frac{1}{r^2} \cdot \left(1 - \frac{1}{g^2(r)}\right) = \varepsilon \cdot \frac{1}{r_0^2}$$

Finalmente, a partir de la representación conforme en E_3 de este tipo de variedad simétrica, que estudiamos en 9.4.6, y pasando ahora a coordenadas cartesianas, el elemento de línea de una variedad 3-D isótropa y por ende homogénea, se puede llevar a la forma canónica:

$$ds^2 = \left(\frac{1}{1 + K \cdot \frac{\left(x_1^2 + x_2^2 + x_3^2\right)}{4}}\right) \cdot \left(dx_1^2 + dx_2^2 + dx_3^2\right) \qquad K = \frac{\varepsilon}{r_0^2}$$

Esta forma, como se verá en el capítulo 12, posee particular importancia en relación con la expresión matemática del postulado de Weyl en el modelo cosmológico estándar.

10. La gravitación y la teoría generalizada de la relatividad

"El mes pasado pasé por uno de los períodos de mi vida más excitantes, absorbentes y, al mismo tiempo más fructíferos períodos de mi vida. No podía pensar en escribir (cartas).
Comprendí que mis anteriores ecuaciones de campo para la gravitación no tenían una base firme…Habiendo perdido toda confianza en la teoría previa, vi claramente que solo se podría alcanzar una solución satisfactoria sobre la base de la teoría covariante general, es decir del (tensor) covariante de Riemann $R_{\mu\nu}$. *Desafortunadamente inmortalicé los últimos errores de este conflicto en mis artículos de la Academia (Prusiana de Ciencias) que pronto enviaré a usted. El resultado final es el siguiente:…Los símbolos de Christoffel* $\Gamma_{\sigma}^{\mu\nu}$ *se deben considerar como la representación natural de los "componentes" del campo gravitatorio…*
La cosa espléndida que experimenté fue que no solo la teoría de Newton se obtiene como una primera aproximación, sino que además, como segunda aproximación se deduce la precesión del perihelio de Mercurio (43''). La magnitud de la deflexión de la luz en el Sol se hace el doble respecto del valor anterior."
Albert Einstein en una carta dirigida a Arnold Sommerfeld en noviembre de 1915.

"El espacio actúa sobre la materia diciéndole cómo moverse. A su vez, la materia reacciona sobre el espacio, diciéndolo como curvarse"
C. Misner, K. Thorne y J. Wheeler "Gravitation" W. Freeman, San Francisco, 1973.

Como la teoría newtoniana de la gravitación resulta incompatible con la relatividad restringida, resulta natural investigar la posibilidad de construir una teoría relativista de la gravitación.
Para compatibilizar la gravitación con la relatividad fue necesario desarrollar una teoría de la gravitación que nos conduce más allá de la teoría restringida y que, como se verá en este capítulo, elimina la fuerza de atracción gravitatoria como causa del movimiento. Esa teoría considera que el movimiento bajo el efecto de la gravedad se puede interpretar como un movimiento libre en un tiempo-espacio curvo cuya curvatura depende de la distribución de la materia en el espacio.
Se apoya en tres principios básicos: equivalencia, covariancia y consistencia.
Su formulación requirió un tipo de espacio abstracto más general que el de Minkowski y el empleo de herramientas matemáticas más generales que las utilizadas en la teoría restringida: las variedades pseudo-riemannianas y el análisis mediante tensores resumidos en el capítulo precedente.

10.1 El principio de equivalencia

En una carta a R. W. Lawson en enero de 1920, Einstein le cuenta que en 1907, mientras trabajaba sobre la forma de incorporar la teoría newtoniana de la gravitación a la teoría de la relatividad, se puso a pensar en las impresiones que experimentaría una

persona que cae desde un tejado. En la medida en que continúe cayendo, para él el campo gravitatorio de la Tierra no existiría: si "dejase caer" objetos de su bolsillo, éstos o bien permanecerían en reposo o en movimiento uniforme respecto del observador en caída libre. Entonces se le ocurrió hacer la siguiente conjetura: durante un período breve de tiempo, las leyes de la física en un pequeño laboratorio en caída libre son las mismas que en ese mismo laboratorio pequeño fijo a un marco inercial en ausencia de campo gravitatorio.
Posteriormente analizó qué ocurriría si al laboratorio se le imponía una aceleración constante. Los objetos "caerían" con la misma aceleración pero en sentido opuesto. Como la masa gravitatoria y la inercial son iguales, desde adentro del laboratorio la situación no sería indistinguible de la de ese mismo laboratorio fijo a un marco inercial en el que existiera un campo gravitatorio cuya aceleración de gravedad tuviera la misma magnitud de la aceleración impuesta pero sentido opuesto.

Así pues, la reflexión sobre la igualdad entre masa gravitacional y masa inercial, condujo a Einstein a plantear la **equivalencia local entre un sistema de referencia no inercial y un campo gravitatorio**.
En una región del espacio lo bastante pequeña como para que allí el campo gravitatorio se pueda considerar uniforme, es posible eliminarlo sustituyéndolo por un sistema uniformemente acelerado cuya aceleración sería igual a la que adquiriría una partícula situada en la región de campo gravitatorio considerada.
Localmente en el tiempo y en el espacio, un sistema acelerado puede aproximarse por uno inercial que, por así decirlo, le es "tangente".

Consideremos todo esto con más detenimiento, comenzando con dos observadores P y P' fijos a un cuerpo que se mueve con un movimiento rectilíneo uniformemente acelerado respecto de un marco inercial K. El segmento de recta que une P con P' es paralelo a la dirección del movimiento. La distancia que separa los puntos mencionados es l_0 y la aceleración que posee el cuerpo (y por ende, que poseen ambos puntos) es a .
Los observadores llevan relojes iguales, sincronizados como se explicó en 6.1. Podría tratarse de relojes de Cesio. Definimos la frecuencia de una oscilación en un determinado sitio, P o P', comparándola con la del reloj local allí situado.
Supongamos que P' marcha delante de P y que P' envía un tren de luz monocromática hacia P, proveniente de su reloj atómico. La frecuencia del tren en P' es $f(P')$.
En el marco de referencia del cuerpo acelerado Ñ, esa señal tarda un intervalo de tiempo $\Delta t = \frac{l_0}{c}$ en ir de P' hasta P.
En el momento en el que la señal partió de P', el cuerpo poseía una velocidad v respecto de K.
Introduzcamos un marco inercial K' que se mueve con esa velocidad respecto de K. Entonces tanto P como P' se encuentran en reposo respecto de K' en el instante en el que se emite el tren. Luego de un intervalo de tiempo Δt la velocidad de Ñ y la de los

puntos P y P' fijos a él, se habrá incrementado en $a \cdot \Delta t = a.\frac{l_0}{c}$. Como consecuencia, cuando es alcanzado por la señal proveniente de P', el punto P poseerá una velocidad $\Delta v = a.\frac{l_0}{c}$ respecto del marco inercial K'.

De acuerdo con lo visto en 8.4.1, el observador en P recibirá la señal con una frecuencia $f(P)$ aumentada por el factor Doppler $\frac{f(P)}{f(P')} = \sqrt{\frac{c+\Delta v}{c-\Delta v}}$.

El reloj ubicado en P' parecerá adelantar (es decir, parecerá oscilar más rápido) respecto del reloj en P.

Si ahora enviamos una señal luminosa de P hacia P', el corrimiento Doppler en la señal será opuesto. La frecuencia recibida en P' vendrá disminuida por un factor respecto de la frecuencia de la señal emitida en P: $\frac{f(P')}{f(P)} = \sqrt{\frac{c-\Delta v}{c+\Delta v}}$

El factor Doppler se puede expresar así en función de la aceleración del cuerpo y de la distancia entre los puntos P y P': $\sqrt{\frac{c+\Delta v}{c-\Delta v}} \approx 1 + \frac{\Delta v}{c} = 1 + \frac{a \cdot l_0}{c^2}$

Supongamos ahora que el cuerpo Ñ es hueco y en su interior se ha hecho vacío. Cualquier objeto que se libere allí dejará de estar acelerado y caerá hacia atrás, es decir en sentido desde P' hacia P. Un observador fijo a la cápsula vería que el objeto presenta una aceleración $-a$ dirigida de P' hacia P, opuesta al sentido del movimiento.

Einstein se dio cuenta de que este movimiento de los objetos podría deberse también a la atracción gravitatoria debida a un planeta sobre el cual se encontrara apoyada la parte posterior de la cápsula[146]. Esto se verificaría siempre y cuando la masa gravitacional y la masa inercial de cada objeto fueran iguales y si la aceleración de la gravedad debida al planeta es $g = -a$ [147].

A partir de este análisis, se puede conjeturar el **principio de equivalencia**: un campo gravitatorio uniforme es equivalente en todos los aspectos a un marco de referencia acelerado Ñ.

Si aplicamos este principio a la cápsula, entre el punto P' y el punto P y si suponemos que es localmente válida la teoría escalar de la gravitación (teoría newtoniana) entonces debe haber una diferencia de potencial gravitatorio

$$\varphi(P') - \varphi(P) = -g \cdot l_0 = a \cdot l_0$$

Como $\frac{f(P)}{f(P')} = \sqrt{\frac{c+\Delta v}{c-\Delta v}}$ y $\sqrt{\frac{c+\Delta v}{c-\Delta v}} \approx 1 + \frac{a \cdot l_0}{c^2}$ resulta que

[146] Y por tanto la totalidad de la cápsula se encontraría en reposo respecto del hipotético planeta.

[147] g se dirige en sentido opuesto a a, por lo cual g es un número negativo.

$$\frac{f(P)}{f(P')} \approx 1 + \frac{(\varphi(P') - \varphi(P))}{c^2} \quad \text{O bien:} \quad \frac{f(P') - f(P)}{f(P')} \approx -\frac{(\varphi(P') - \varphi(P))}{c^2}$$

Consideremos dos puntos con diferente potencial gravitatorio. Entonces un átomo excitado situado en el punto donde el potencial gravitatorio es menor emite líneas espectrales desplazadas al rojo respecto de un átomo de la misma clase pero situado en el otro punto. Esto suministra un fundamento, basado en el principio de equivalencia, a las consideraciones efectuadas en 8.3.

De acuerdo con la teoría restringida de la relatividad, no habría motivo para suponer que dos relojes idénticos en reposo relativo puedan presentar una diferencia en su período de oscilación. Ahora vemos que esa diferencia debe esperarse cuando ambos relojes se encuentran fijos a un cuerpo acelerado o cuando ambos se encuentran a diferente potencial gravitatorio.

La proporcionalidad (que puede transformarse en una igualdad utilizando unidades adecuadas) entre la masa inercial y la masa gravitatoria es sorprendente: a priori parece que no hubiera una razón por la cual tendrían que estar relacionadas.

Al fin y al cabo, la masa inercial es el coeficiente de resistencia a la aceleración de un cuerpo en la formulación matemática de las leyes de la dinámica, mientras que la masa gravitatoria es un análogo de la carga en el caso eléctrico, puesto que es el coeficiente de acoplamiento del cuerpo al campo gravitatorio[148].

Mientras que el movimiento de una partícula cargada en un campo electromagnético depende de dos parámetros independientes, su masa inercial y su carga eléctrica (que es su coeficiente de acoplamiento al campo), en un campo gravitatorio la igualdad entre la masa inercial y la masa gravitatoria elimina la masa del cuerpo de las ecuaciones del movimiento: dos cuerpos con las mismas condiciones iniciales se mueven igual, sea cual sea la diferencia entre sus masas.

El principio de equivalencia puede formularse en dos versiones.

Una débil, planteada por Einstein en 1907, suficiente para plantear una teoría de la gravitación:

"En el interior de una región lo bastante localizada del tiempo-espacio, adyacente a una concentración de masas, el movimiento de los cuerpos sometidos exclusivamente a efectos gravitatorios no puede ser distinguido por vía de un experimento, cualquiera que sea, del movimiento de los cuerpos en el interior de una región con una aceleración uniforme adecuada". Se sobre-entiende que se trata de lo que en física se denominan cuerpos de prueba, o sea en este caso cuerpos tales que no modifican significativamente el campo gravitatorio del lugar donde se encuentran.

---

[148] Es posible distinguir dos masas gravitatorias asignables a un cuerpo: la masa considerada como **carga pasiva** $m_{g,cp}$ que determina la fuerza que un campo gravitatorio ejerce sobre ese cuerpo, y la masa considerada como **carga activa** $m_{g,ca}$ que determina el campo gravitatorio generado por el cuerpo. La igualdad $m_{g,cp} = m_{g,ca}$ de ambas masas gravitatorias es necesaria para la igualdad de las fuerzas gravitatorias de interacción entre dos cuerpos. Sobre esta distinción y sus consecuencias ver D. Ivanenko y G. Sardanashvili "Gravitación", capítulo 2, Editorial URSS, Moscú, 2005.

Una fuerte, planteado posteriormente y que abarca, además de las interacciones gravitatorias, las interacciones electromagnéticas y las interacciones nucleares:

"En el interior de una región lo bastante localizada del tiempo-espacio, adyacente a una concentración de masas, el comportamiento físico de los cuerpos no puede ser distinguido por vía de un experimento, cualquiera que sea, del comportamiento físico de los cuerpos en el interior de una región con una aceleración uniforme adecuada".

La versión fuerte implica la débil, pero la débil podría permanecer válida aunque la versión fuerte debiera ser descartada.

Algunas consecuencias de ambas versiones han sido confrontadas con resultados experimentales y hasta el momento parecen bien confirmadas.

Las confirmaciones más precisas de la forma débil (a menos de una parte en 10^{12}) se obtuvieron en experimentos destinados a comparar las aceleraciones de la Tierra y de la Luna en el campo de gravitación del Sol, efectuados a fines del siglo XX y comienzos del actual. Se espera una confirmación con una precisión de una parte en 10^{18} cuando se pueda llevar a cabo una misión espacial diseñada a tal efecto (Satellite Test of the Equivalence Principle, STEP).

Los experimentos para confrontar la forma fuerte son más difíciles de llevar a cabo y en general resultan más ambiguos: se focalizan en determinar una posible variación de la constante gravitatoria G con el paso del tiempo, sobre la base de que la mayor parte de las teorías de la gravitación alternativas a la de Einstein, que violan la forma fuerte del principio predicen valores locales de dicha constante variables. La constancia de G parece confirmada a menos de una parte en 10^{13} por año.

10.2 El tiempo-espacio-materia y la gravitación

La mecánica clásica separa el espacio del tiempo. Asume la homogeneidad e isotropía del espacio, asimilándolo al modelo euclidiano. Define invariantes bajo traslaciones y rotaciones en el espacio tridimensional que conducen a las leyes de conservación de la cantidad de movimiento lineal y angular en los sistemas mecánicos. Por otro lado describe la homogeneidad del tiempo y describe invariantes bajo una traslación en el tiempo que conducen a la ley de conservación de la energía.

La conexión entre espacio y tiempo se realiza en las transformaciones de Galileo: el tiempo permanece incambiado al pasar de un marco de referencia inercial a otro, pero las coordenadas espaciales se modifican cuando los marcos se hallan en movimiento relativo. Interesan las propiedades que permanecen invariantes bajo el grupo de transformaciones de Galileo, porque corresponden a los mismos procesos descriptos por diferentes observadores ubicados en distintos marcos inerciales: el aspecto objetivo de la realidad según la física newtoniana.

La teoría restringida de la relatividad introduce un continuo tiempo-espacio tetra-dimensional, el espacio no euclidiano de Minkowski, con una pseudo-métrica definida mediante el denominado intervalo entre eventos. Asume la homogeneidad e isotropía del espacio de Minkowski, lo que se manifiesta en la invariancia del intervalo bajo el grupo de transformaciones de Lorentz.

Interesan las propiedades que permanecen invariantes bajo el grupo de transformaciones de Lorentz, porque corresponden a los mismos procesos descriptos por diferentes observadores situados en diferentes marcos inerciales: el nuevo aspecto objetivo de la realidad según la física relativista, que sustituye a la objetividad asumida en la física newtoniana.

La teoría generalizada de la relatividad elimina la restricción a los marcos inerciales. Los marcos de referencia pueden ser cualesquiera en movimiento relativo arbitrario.
Lo único que se asume es que todo evento se puede describir mediante una cuaterna de coordenadas $\left(x^0, x^1, x^2, x^3\right)$ (números reales) respecto de un marco K y que sus coordenadas respecto de otro marco K' vienen dadas por transformaciones puntuales regulares e invertibles $x'^j = f^j\left(x^0, x^1, x^2, x^3\right)$ para j=0, 1, 2, 3. En principio, aunque en la práctica no se haga, un marco de referencia podría consistir en cuatro aviones volando de cualquier forma. Las cuatro coordenadas del evento podrían ser los instantes de tiempo en los que una señal procedente del evento es registrada por cada uno de los cuatro aviones. Ese tiempo podría ser registrado por un reloj no newtoniano (ver 2.1).
Pero en la formulación de la teoría de la gravitación, tal como se lleva a cabo en la práctica, se utilizan tres coordenadas asociadas con el espacio físico x^1, x^2, x^3 y una coordenada x^0 asociada con el tiempo.
Ahora interesan en principio las propiedades que permanecen invariantes bajo el grupo formado por todas las transformaciones puntuales regulares e invertibles: $x'^j = f^j\left(x^0, x^1, x^2, x^3\right)$ j=0, 1, 2, 3.
En consecuencia **todos los marcos de referencia** se deberían poder utilizar en una descripción de los eventos del mundo físico, ya no solo los marcos inerciales.
Aparece un nuevo aspecto objetivo de la realidad según la relatividad generalizada, que debe sustituir a la objetividad asumida en la teoría restringida, como la objetividad de esta última sustituyó a la objetividad de la física newtoniana.

Con la generalización del principio de relatividad a marcos de referencia arbitrarios parece que se puede llevar a cabo el programa planteado por Mach: eliminar el último vestigio de carácter absoluto en el tiempo y en el espacio, transformándolos en meros medios para la descripción de los fenómenos físicos.
Pero mientras que Mach insistió siempre en los aspectos relativos del espacio y el tiempo[149], Einstein se dispuso a encontrar las relaciones espacio-temporales que permanecían invariantes bajo el grupo de todas las transformaciones puntuales: salió a la caza de nuevos absolutos.

---

[149] La importancia que Ernst Mach dio a la relatividad del espacio y del tiempo le impidió dar una respuesta positiva al programa que el mismo había planteado. Siempre se opuso a la respuesta positiva de Einstein frente a ese programa, es decir, a la teoría de la relatividad.
Resulta interesante que, por motivos muy similares, Mach también se oponía a las teorías que postulaban átomos que no eran observables. Ya hacia el final de sus días, algunas personas le mostraron los fenómenos que se interpretan como impactos (sobre pantallas fluorescentes) de partículas provenientes de desintegraciones radioactivas.

Para cumplir este programa se pueden utilizar como guía dos principios: el principio general de covariancia y el principio general de consistencia.

10.2.1 Covariancia

El principio general de covariancia es la máxima extensión posible del principio de relatividad: las leyes de la física deben adoptar la misma forma en todos los marcos de referencia.
Einstein concretó este principio en un requerimiento sobre las ecuaciones que expresan leyes físicas: deben mantener su forma bajo una amplia clase de transformaciones de coordenadas de tiempo y espacio.
Pudo satisfacer este requerimiento de invariancia (denominado **covariancia generalizada**) recurriendo a formular las ecuaciones de la física en forma de igualdades entre tensores involucrando las diversas operaciones algebraicas, contracción de índices y derivación covariante, resumidos en la parte 9.5.

No obstante, Kretschmann (un profesor alemán de enseñanza media) señaló en un artículo publicado en 1917, que esto por sí solo no basta para asegurar que una teoría verifique un principio de relatividad[150]. Casi cualquier teoría puede matematizarse de modo tal que resulte covariante bajo el grupo de transformaciones asociado con el principio de relatividad: hace falta algo más. En el caso de la teoría de la gravitación de Einstein ese algo más es la idea de generalizar el intervalo entre eventos de la teoría restringida asumiendo que las componentes del intervalo dependen de la distribución de materia: esta es una concreción clave en torno a la cual se puede implementar el requerimiento de covariancia.

Para expresar matemáticamente la idea de que el movimiento debido exclusivamente a los efectos de la gravitación es un movimiento libre en un espacio curvo cuya curvatura depende de la distribución de la masa-energía, es necesario comenzar por la variedad que corresponde al tiempo-espacio.
En particular, el intervalo ds^2, si es una generalización del invariante que se encuentra en la base de la teoría restringida, debe corresponder a una pseudo-métrica en la variedad formada por el tiempo-espacio-materia. Entonces debe permanecer invariante bajo la clase de transformaciones de coordenadas que presumiblemente conectan las coordenadas de un mismo evento cuando es observado desde diferentes marcos de referencia.
La expresión local del intervalo que corresponde a un continuo tiempo-espacio se expresa así: $ds^2 = \sum_{i,k=0}^{3} g_{i,k} \cdot dx^i \cdot dx^k$ donde los índices i,k se suman sobre los números naturales entre 0 y 3, representando con 0 la coordenada temporal en unidades de longitud $x^0 = c \cdot t$ y con los demás números las coordenadas espaciales. Los símbolos $g_{i,k}$ varían por lo general de un punto a otro del tiempo-espacio y forman el tensor métrico que determinan la pseudo-métrica local de una variedad tetra-dimensional.
Para que ds^2 posea dimensiones de cuadrado de una longitud, las componentes $g_{i,k}$ del tensor métrico deben ser no dimensionales. La forma de variación de estas componentes, de un punto a otro, determina el valor local del tensor de curvatura del

[150] Ver J. Norton, "The physical content of general covariance" en J. Eiseinstaedt y A. Kox (Eds.) "*Studies in the history of general relativity*", Birkhäuser, Boston, 1999, pp. 281-315.

continuo tiempo-espacio-materia[151] y también las ecuaciones de las geodésicas (que corresponden al movimiento libre), como se explicó en 9.5.
En consecuencia, el primer objetivo de la teoría debe ser hallar ecuaciones que permitan determinar las componentes del campo tensorial $g_{i,k}$ en función de la distribución de la materia. Para hallarlas se dispone de una guía que suministra el principio de consistencia.

10.2.2 Consistencia y tanteos preliminares

El principio de consistencia, como vimos a propósito del caso de la teoría restringida y la mecánica newtoniana al comienzo de 8.1, puede resumirse en el siguiente requerimiento: toda teoría que aspire a sustituir a otra debe dar cuenta de las predicciones exitosas de las teorías previas.
En el caso de la teoría generalizada, esto implica que debe ser consistente tanto con la teoría restringida como con la gravitación newtoniana cuando estas últimas resulten aplicables.

La consistencia con la teoría restringida impone un primer tipo de restricción sobre la naturaleza del continuo tiempo-espacio-materia. Esta restricción se puede hallar a partir de la idea de que los efectos gravitatorios equivalen a una curvatura de ese continuo.
Si la curvatura desaparece, desaparecen esos efectos y deberíamos esperar que al enderezarse el tiempo-espacio-materia se transforme en un tiempo-espacio de Minkowski.
Pero es posible avanzar más en esta dirección. Considerando al tiempo-espacio-materia como una variedad regular curvada cuyos puntos son eventos, en cada punto P del se define una variedad lineal tangente.
La variedad tangente posee el mismo número de dimensiones (4) que la variedad curvada, pero su curvatura es en todas partes nula.
No obstante en un entorno lo bastante pequeño de un evento P_0, la variedad de Riemann se puede aproximar por su variedad tangente con un error que es un infinitésimo de orden superior al de las diferencias de coordenadas de tiempo y espacio entre los puntos P de ese entorno y P_0.
Cuando se fijan las componentes del tensor métrico en el valor $g_{i,k}(P_0)$ que poseen en P_0, el elemento de línea resultante $ds^2 = \sum_{i,k=0}^{3} g_{i,k}(P_0) \cdot dx^i \cdot dx^k$ no es otra cosa que la pseudo-métrica euclidiana sobre la variedad lineal tangente. Esta es plana y deberíamos esperar que fuera un espacio de Minkowski.
Así pues la consistencia con la relatividad restringida exige que la variedad que representa el continuo tiempo-espacio-materia se pueda aproximar localmente en cada uno de sus puntos por un tiempo-espacio de Minkowski tangente, y se transforme en un único tiempo-espacio de Minkowski si los efectos gravitatorios desaparecen en todas partes.

---

[151] Como se puso en evidencia a lo largo del capítulo 9, para medir la curvatura local de una variedad es suficiente efectuar mediciones desde adentro. No es necesario imaginar un espacio euclidiano de mayor dimensión en el interior del cual ese continuo se encuentra inmerso, en forma análoga a como una superficie curvada bidimensional se encuentra inmersa en el espacio euclidiano tridimensional que nos es familiar. Pero si se prefiere imaginar el mundo tetra-dimensional curvado y embebido en un espacio euclidiano, éste debe tener 10 dimensiones de acuerdo con la fórmula vista en la sección 9.4.4.

Por otra parte las ecuaciones que en la teoría restringida expresan leyes físicas como las del electromagnetismo, que no se reducen a efectos gravitatorios, deben constituir aproximaciones locales a las correspondientes ecuaciones de la teoría generalizada.
Desde el punto de vista físico todo esto se puede resumir diciendo que todos los resultados de la teoría restringida se deberían verificar como aproximaciones locales en la teoría generalizada.
Para ver qué tipo de restricción se desprende de la consistencia con la teoría de la gravitación de Newton, es necesario replantear esta teoría en términos de campos.
Este replanteo se llevó a cabo en lo fundamental durante el siglo XIX, y se puede resumir en una relación entre el potencial gravitatorio $\varphi(t,\vec{r})$ y la densidad de materia $\rho(t,\vec{r})$ de la teoría newtoniana en la forma de ecuación a derivadas parciales, debida a Poisson:

$$\frac{\partial^2\varphi(t,\vec{r})}{\partial x^2}+\frac{\partial^2\varphi(t,\vec{r})}{\partial y^2}+\frac{\partial^2\varphi(t,\vec{r})}{\partial z^2}=4\pi\cdot G\cdot\rho(t,\vec{r})$$

El operador Δ (el operador de Laplace) está formado por las derivadas parciales de segundo orden respecto de las coordenadas espaciales y G es la constante de gravitación de Newton.
La densidad de masa constituye la fuente del campo gravitatorio, que en cada punto y en cada instante viene dado por el vector $\vec{g}(t,\vec{r})=-\nabla\varphi(t,\vec{r})$ de componentes cartesianas $-\frac{\partial\varphi}{\partial x}$, $-\frac{\partial\varphi}{\partial y}$ y $-\frac{\partial\varphi}{\partial z}$.
Como, a diferencia de lo que acontece en la ecuación de ondas clásica, en la ecuación de Poisson solo aparecen derivadas respecto a las coordenadas espaciales, pero no respecto del tiempo, la propagación del efecto de una variación temporal en la densidad sobre el potencial, y por ende sobre el campo gravitatorio, se produce en forma instantánea.
Ya en sus primeros trabajos de 1904, 1905 y 1906 sobre la teoría restringida, Poincaré insistió en la universalidad del principio de relatividad y propuso su extensión a las interacciones gravitatorias. Asumiendo una velocidad finita de propagación de las interacciones gravitatorias, igual a la velocidad de la luz, introdujo un retardo en la formulación de la ley de atracción. Esta forma de abordar el problema no resultó adecuada.
En el período comprendido entre la aparición de los trabajos pioneros de Poincaré y los trabajos fundacionales de Hilbert y de Einstein en 1915 y 1916 sobre las ecuaciones de la teoría generalizada, varios investigadores, entre ellos Minkowski, Abraham, Nordström y Einstein mismo, intentaron construir teorías escalares de la gravitación.
En líneas generales estas teorías buscaban una generalización de la ecuación de Poisson[152] $\Delta\varphi=4\pi\cdot G\cdot\rho$ en el marco de la relatividad restringida y aplicando el principio de consistencia.
Si se supone que el potencial escalar y la densidad de masa son invariantes relativistas, una generalización inmediata que resulta compatible con el principio de relatividad restringido es la siguiente ecuación de ondas clásica para el potencial gravitatorio[153]:

---

[152] En coordenadas cartesianas el operador de Laplace Δ se expresa así: $\Delta\varphi=\frac{\partial^2\varphi}{\partial x^2}+\frac{\partial^2\varphi}{\partial y^2}+\frac{\partial^2\varphi}{\partial z^2}$

$$\frac{\partial^2\varphi}{\partial x^2}+\frac{\partial^2\varphi}{\partial y^2}+\frac{\partial^2\varphi}{\partial z^2}-\frac{1}{c^2}\cdot\frac{\partial^2\varphi}{\partial t^2}=4\pi\cdot G\cdot\rho_0$$

En el término $4\pi\cdot G\cdot\rho_0$, fuente del campo escalar $\varphi(t;x,y,z)$, aparece la densidad de masa en reposo $\rho_0(t;x,y,z)$.

No obstante todos los esfuerzos realizados, cuando se confrontaron con los hechos físicos, las teorías escalares fracasaron.

La masa relativista total, y no solo la masa en reposo, parecían ser la fuente de las interacciones gravitatorias, de modo que las contribuciones asociadas con la velocidad de movimiento de la materia debían ser tenidas en cuenta.

Así el siguiente paso consistió en intentar introducir un cuadripotencial gravitatorio cuyas fuentes se encontraran en el cuadrivector energía-cantidad de movimiento.

Pero los esfuerzos realizados en este sentido también fracasaron, entre otros motivos debido a que el término fuente de los efectos de gravitación no se podía representar como un cuadrivector. Todo apuntaba a pensar que se requerían objetos matemáticos más complejos, tanto para representar las fuentes del campo como para representar el campo mismo.

La caracterización del movimiento, por lo demás libre, de una partícula newtoniana restringida a moverse en una superficie, estudiado en la parte 9.3, parecía sugerir una salida de este impasse. Este fue el camino finalmente seguido por Einstein y por Hilbert para formular las ecuaciones de una teoría relativista de la gravitación, utilizando el tensor métrico como expresión del campo gravitatorio, el tensor de energía impulsión como expresión de las fuentes del campo y las ecuaciones de las geodésicas como ecuaciones del movimiento por efecto de la gravedad.

10.2.3 Las ecuaciones del campo gravitatorio

Para formular las ecuaciones del campo gravitatorio en presencia de materia-energía (partículas masivas y radiación) Einstein comenzó de nuevo con la ecuación de Poisson de la teoría newtoniana de la gravitación: $\Delta\varphi=4\pi\cdot G\cdot\rho$

Utilizándola como guía, sustituyó la densidad de masa ρ por una versión generalizada del tensor de energía-impulsión $T^{i,k}$ $(i,k=0,1,2,3)$ visto ya (en una versión restringida) en la parte 8.7 a propósito de la dinámica de fluidos. Es un tensor simétrico de orden dos que describe la distribución y el flujo de energía y cantidad de movimiento de la materia y de la radiación en un lugar del continuo tiempo-espacio-materia.

Este tensor se debería relacionar con el campo gravitatorio de tal manera que se asegurara la validez del principio de conservación de la energía-cantidad de movimiento. En la teoría restringida de la relatividad este principio único sustituye a los tres principios fundamentales de conservación de la masa, la energía y la cantidad de movimiento de la física clásica[154].

---

[153] Esta ecuación conduce a predecir la existencia de ondas gravitacionales.

[154] En la parte 8.7 vimos su aplicación a la mecánica de un fluido ideal en el tiempo-espacio de Minkowski. En este caso, en el cual la variedad se plana, la divergencia covariante se reduce a la cuadridivergencia vista en 8.7.

Como vimos al final de la sección 9.4.2, efectuando una contracción de índices a partir de la derivada covariante de un tensor escrito en forma contravariante obtenemos el tensor denominado **divergencia covariante**.
La expresión matemática que generaliza el principio de conservación de la energía-cantidad de movimiento en una variedad riemanniana es la anulación de la divergencia covariante del tensor de energía-impulsión: $\sum_{l=0}^{3} D_l T^{i,l} = 0$
En su búsqueda de la expresión que estaría destinada a sustituir a $\Delta\varphi$, Einstein tuvo en cuenta que para tener el mismo orden y simetría que el tensor de energía-impulsión debería ser un tensor de orden dos, simétrico.
Para poder expresar la identidad presumida entre los efectos gravitatorios y la curvatura de la variedad tiempo-espacio, ese tensor destinado a sustituir a $\Delta\varphi$ debería estar construido a partir del tensor métrico.
El tensor de curvatura depende de los coeficientes de conexión de Riemann-Christoffel como se mostró en 9.5.4: $R^{l}{}_{i,k,j} = \frac{\partial \Gamma^{l}{}_{i,k}}{\partial x^{j}} - \frac{\partial \Gamma^{l}{}_{i,j}}{\partial x^{k}} + \sum_{m=0}^{3} \Gamma^{l}{}_{m,j} \cdot \Gamma^{m}{}_{i,k} - \sum_{m=0}^{3} \Gamma^{l}{}_{m,k} \cdot \Gamma^{m}{}_{i,j}$
Los coeficientes de conexión se expresan a su vez en términos de las componentes del tensor métrico y de sus derivadas parciales respecto de la coordenada temporal y las tres coordenadas espaciales: $\Gamma^{i}{}_{j,k} = \frac{1}{2} \cdot \sum_{l=1}^{n} g^{i,l} \cdot \left(\frac{\partial g_{l,k}}{\partial x^{j}} + \frac{\partial g_{j,l}}{\partial x^{k}} - \frac{\partial g_{j,k}}{\partial x^{l}} \right)$ Como desaparecen si las componentes del tensor métrico no varían de un punto a otro, Einstein conjeturó que en el tensor que se buscaba deberían aparecer las derivadas parciales de esas componentes respecto de las coordenadas. Cabía esperar que aparecieran derivadas parciales de componentes $g_{i,k}$ hasta el segundo orden, puesto que hasta ese orden aparecen en la descripción de la curvatura de las variedades. En ese caso el tensor métrico ocuparía en la nueva teoría de la gravitación el lugar que ocupa del potencial newtoniano en la teoría antigua. Como afirma Einstein en la carta a Sommerfeld que aparece al comienzo de este capítulo, los coeficientes $\Gamma^{i}{}_{j,k}$ ocuparían a su vez el lugar del vector que representa al campo gravitatorio en la teoría de Newton.

Por analogía con $\Delta\varphi$, supuso que el tensor que se buscaba sería lineal y homogéneo en términos de las derivadas parciales de segundo orden.
Además, tuvo en cuenta que la divergencia covariante correspondiente al término tensorial destinado a sustituir a $\Delta\varphi$ en el lado izquierdo de la ecuación, debería anularse[155].
A partir de todos estos requisitos conjeturó que en el miembro de la izquierda de la ecuación se podría encontrar el tensor de Ricci[156] $R^{i,k}$ y el tensor métrico $g^{i,k}$ en una

[155] Ver, por ejemplo, el capítulo 11 (ecuaciones del campo de gravitación) del libro de L. Landau y E. Lifchitz "Teoría del campo", Reverté, Barcelona, 1982.

[156] El tensor de Ricci en su forma dos veces covariante $R_{i,k}$ se obtiene por contracción de índices en el tensor de curvatura de la variedad de Riemann que representa al continuo tiempo-espacio y posterior. La forma dos veces contravariante $R^{i,k}$ se obtiene por elevación de índices empleando la forma contravariante $g^{i,k}$ del tensor métrico, como se explicó en la sección 9.5.1.

combinación tal que su divergencia covariante resultara ser siempre nula.
La combinación más simple que verifica todas estas restricciones es el tensor de Einstein $G^{i,k} = R^{i,k} - \frac{1}{2} \cdot g^{i,k} \cdot R$

En esta expresión el coeficiente $\frac{1}{2}$ asegura que se verifique la condición de divergencia covariante nula: $\sum_{l=0}^{3} D_l G^{i,l} = 0$ y la curvatura escalar R se define a partir del tensor de Ricci como se mostró en 9.5.4: $R = \sum_{i,k=1}^{n} g^{i,k} \cdot R_{i,k}$

Finalmente, Einstein propuso las siguientes ecuaciones para el campo gravitatorio:

$$G^{i,k} = R^{i,k} - \frac{1}{2} \cdot g^{i,k} \cdot R = -\kappa \cdot T^{i,k} \qquad i,k = 0,1,2,3$$

La constante, $\kappa = \frac{8\pi}{c^4} \cdot G$ es proporcional a la constante de gravitación de Newton G. Conocida como constante de gravitación de Einstein, multiplica al tensor $T^{i,k}$ y garantiza la consistencia con la teoría newtoniana de la gravitación.

Como el tensor de curvatura de Ricci se forma a partir de las componentes del tensor métrico y de sus derivadas, el miembro de la izquierda de la ecuación, es decir, el tensor de Einstein $G^{i,k}$, depende solamente del tensor que representa las propiedades métricas del tiempo-espacio.
El miembro de la derecha presenta el tensor $T^{i,k}$ asociado, como ya dijimos, a las propiedades de la distribución de la materia (incluyendo el campo electromagnético) que constituyen las fuentes del campo gravitatorio: "representa la energía que engendra el campo gravitatorio; pero es en sí misma, de carácter no gravitatorio, como sucede por ejemplo con la energía del campo electromagnético, de la densidad de la materia ponderable, etc. Para representar este tensor $T_{i,k}$ se utilizan conceptos de Física pre-relativista que sólo a posteriori se han adaptado al principio de relatividad general." [157]

Estas ecuaciones de campo, junto con condiciones de borde adecuadas, permiten determinar las componentes del tensor métrico en cada punto del continuo tiempo-espacio, suponiendo que se supone conocido el tensor energía-impulsión.
La simetría de los tensores involucrados en las ecuaciones de campo tiene como consecuencia una disminución en el número de ecuaciones independientes de 16 a 10, lo cual permite hallar las 10 componentes independientes del tensor simétrico en su forma covariante $g_{i,k}$ o, en su forma contravariante equivalente $g^{i,k}$.
Para que una región de una variedad pueda considerarse plana, deben anularse las componentes del tensor de curvatura $R^{l}{}_{i,k,j}$ en todos sus puntos. Esto implica que el tensor de Ricci $R^{i,k}$ se anula en esos puntos. Pero debido a como se lo define

[157] A. Einstein "*El significado de la relatividad*", 2ª edición, Espasa Calpe Argentina, Buenos Aires, 1952, Apéndice II.

($R_{i,k} = \sum_{l=1}^{n} R^{l}{}_{i,k,l}$) de la anulación del tensor de Ricci no se desprende que el tiempo-espacio sea plano, o sea que $R^{l}{}_{i,k,j}$ sea nulo. Si el tensor de Ricci es nulo, entonces debe anularse el tensor de energía-impulsión en ese punto: hay vacío en lugar de radiación o partículas.

Aplicando la operación de descenso de índices de un tensor, la ecuación de Einstein para el campo gravitatorio se puede escribir también con todos los tensores en forma covariante: $R_{i,k} - \frac{1}{2} \cdot g_{i,k} \cdot R = -\kappa \cdot T_{i,k} \qquad i,k = 0,1,2,3$

Es posible demostrar[158] que la ecuación $R_{i,k} - \frac{1}{2} \cdot g_{i,k} \cdot R = -\kappa \cdot T_{i,k}$ es equivalente a la ecuación $R_{i,k} = -\kappa \cdot \left(T_{i,k} - \frac{1}{2} \cdot g_{i,k} \cdot T \right)$

Por definición, el escalar $T = \sum_{l=0}^{3} T^{l}{}_{l}$ se construye por contracción de índices a partir de la forma mixta $T^{i}{}_{j} = \sum_{k=0}^{3} g^{i,k} \cdot T_{k,j}$ del tensor de energía-impulsión.

De $R_{i,k} = -\kappa \cdot \left(T_{i,k} - \frac{1}{2} \cdot g_{i,k} \cdot T \right)$ se desprende que si en un punto se anula el tensor de energía-impulsión, o sea si en ese punto hay vacío, entonces en ese punto también se anula el tensor de Ricci. Por tanto la anulación del tensor de Ricci en un punto equivale a que en ese punto hay vacío, y justifica la denominación "soluciones en el vacío" con que se conocen las soluciones de la ecuación tensorial $R_{i,k} = 0$ (o de su ecuación equivalente $R^{i,k} = 0$). En una aproximación de campo gravitatorio débil, se obtienen soluciones de vacío en forma de ondas de gravitación que viajan con la velocidad de la luz y que pueden estar localizadas en regiones acotadas del espacio físico. La teoría generalizada muestra que esas ondas se pueden generar por radiación a partir de las masas en movimiento de un sistema de cuerpos materiales, localizado en una región acotada. La pérdida de energía que esta radiación conlleva es lo bastante pequeña como para que su efecto sobre el movimiento de las masas no sea significativo, al parecer ni siquiera en el caso de objetos a escala astronómica ni a una escala de tiempo cósmica[159].

---

[158] La demostración puede verse, entre otras fuentes, en R. Lambourne, "*Relativity, gravitation and cosmology*", p. 294, Cambridge University Press, U.K. 2010.

[159] No obstante, se ha podido obtener evidencia indirecta sobre las ondas de gravitación. Esta evidencia se basa en haber observado durante 30 años la disminución del período orbital de un pulsar binario conocido como PSR B1913+16. Al parecer estaría formado por dos estrellas de neutrones (las estrellas de neutrones se consideran más adelante en la sección 13.3.1), una de las cuales gira sobre sí misma unas 17 veces por segundo y emite pulsos regulares de radiación electromagnética que se pueden detectar en nuestro planeta. Las modificaciones de esta señal permiten inferir la variación en el período de rotación, estimado en 7.75 horas para estrellas de masas próximas a 1.4 veces la masa solar. El desacuerdo entre lo observado y lo que predice la teoría generalizada suponiendo que el efecto se debe en última instancia a la radiación de ondas de gravitación es de solo un 0.2 %. Sobre ondas gravitacionales y su evidencia observacional, ver por ejemplo, R. Lambourne, "*Relativity, gravitation and cosmology*", pp. 226-232, Cambridge University Press, U.K. 2010.

10.2.4 La ley del movimiento en un campo gravitatorio

En la formulación de 1915 y 1916 de la teoría de la gravitación, Einstein agregó un postulado adicional: "Los cuerpos, en presencia solo de un campo gravitatorio, se mueven a lo largo de geodésicas en el tiempo-espacio curvado". Se reduce en última instancia la gravitación a la curvatura del espacio.

Las líneas geodésicas en el tiempo-espacio-materia se pueden clasificar en tres tipos:

El **tipo temporaloide** caracterizado porque el intervalo $ds^2 = \sum_{i,k=0}^{3} g_{i,k} \cdot dx^i \cdot dx^k$ es siempre positivo a lo largo de la curva.

El **tipo espacialoide** caracterizado porque el intervalo $ds^2 = \sum_{i,k=0}^{3} g_{i,k} \cdot dx^i \cdot dx^k$ es siempre negativo a lo largo de la curva.

El **tipo nulo** caracterizado porque el intervalo $ds^2 = \sum_{i,k=0}^{3} g_{i,k} \cdot dx^i \cdot dx^k$ es siempre nulo a lo largo de la curva.

El postulado introducido por Einstein se puede concretar un poco más en un principio de movimiento geodésico:
"Las líneas geodésicas temporaloides representan líneas de universo posibles para partículas con masa en reposo no nula que caen libremente bajo la sola influencia de la gravedad. Las geodésicas nulas representan líneas de universo posibles para partículas cuya masa en reposo es nula y que se mueven bajo la sola influencia de la gravitación."

Así pues, bajo el solo efecto del campo gravitatorio un planeta orbita en torno al sol siguiendo una geodésica temporaloide en el continuo tiempo-espacio curvado adyacente a la estrella, y un rayo de luz moviéndose en las proximidades del sol sigue una geodésica nula en ese mismo tiempo-espacio.

La solución de las ecuaciones locales del campo gravitatorio $G^{i,k} = -\kappa \cdot T^{i,k}$ $(i,k = 0,1,2,3)$, en una región tetra-dimensional del tiempo-espacio y que presenta interés desde el punto de vista físico, cuando se supone conocido el tensor de energía-impulsión en el interior de esa región y se añaden condiciones en su frontera adecuadas, permite determinar las componentes del tensor métrico.
A su vez, a partir de estas componentes se calculan los coeficientes de conexión, las posibles líneas de universo de la radiación y de los cuerpos materiales en el interior de esa región, así como el tensor de curvatura de Riemann-Christoffel y otras propiedades de interés del continuo tiempo-espacio-materia.

Las geodésicas, sea cual sea su tipo en el tiempo-espacio curvo, se determinan a partir de los coeficientes $g_{i,k}$, generalmente variables de un punto a otro del espacio, a partir de las ecuaciones presentadas en 9.5: $\frac{d^2 x^i(u)}{du^2} + \sum_{j,k} \Gamma_{j,k}{}^{i} \cdot \frac{dx^j(u)}{du} \cdot \frac{dx^k(u)}{du} = 0$

Investigaciones posteriores[160] sobre la ley de conservación $\sum_{l=0}^{3} D_l T^{i,l} = 0$ ($i = 0,1,2,3$) y su relación con las ecuaciones del campo gravitatorio $G^{i,k} = -\kappa \cdot T^{i,k}$ ($i,k = 0,1,2,3$) mostraron que la condición de anulación de la divergencia covariante del tensor de energía impulsión implica el movimiento geodésico de la radiación y los cuerpos materiales.

10.3 La relatividad restringida y la gravitación newtoniana como casos límite. Deducción de la conexión entre κ y G. [161]

Las componentes (siempre no dimensionales) del tensor métrico se pueden expresar así: $g_{i,k} = m_{i,k} + h_{i,k}$ Aquí $m_{i,k}$ son las componentes del tensor métrico para un tiempo-espacio de Minkowski tetra-dimensional y $h_{i,k}$ es una perturbación asociada con la curvatura del continuo tiempo-espacio-materia[162].
Ahora bien, hay esencialmente un espacio de Minkowski.
La diferencia en los tensores $m_{i,k}$ de un punto a otro del continuo tiempo-espacio-materia se debe solamente a las coordenadas utilizadas para describir la variedad.
Pero en todos los casos un cambio de coordenadas en cada variedad de Minkowski tangente lleva el tensor métrico de esa variedad a la forma canónica $m_{0,0} = +1$, $m_{1,1} = m_{2,2} = m_{3,3} = -1$ y $m_{i,k} = 0$ siempre que $i \neq k$.
Si las componentes del tensor de curvatura de Riemann-Christoffel se aproximan a cero en todos los puntos (eventos) de la variedad, las derivadas parciales de las perturbaciones $h_{i,k}$ se aproximan a cero en todos los puntos del tiempo-espacio-materia. Entonces la variedad riemanniana de la teoría generalizada se aproxima al tiempo-espacio plano de la teoría restringida y debido a como se definieron las perturbaciones, éstas deben tender a cero. Y si las $h_{i,k}$ tienden a cero, las componentes del tensor de curvatura tienden a cero y la variedad límite es plana.

Supongamos que, además de ser pequeñas respecto de 1 las magnitudes de las perturbaciones $h_{i,k}$, las velocidades del movimiento de los cuerpos materiales es en todas partes pequeña respecto de la velocidad de la luz y dichos cuerpos no interactúan

---

[160] W. Rindler "*Relativity: special, general and cosmological*", Oxford University Press, U.K. 2001.
M. Hobson, G. Efstathiou y A. Lasenby, "*General relativity: an introduction for physicists*", Cambridge University Press, Cambridge, U.K, 2006.
C. Misner, K. Thorne y J. Wheeler "*Gravitation*" W. Freeman, San Francisco, 1973.

[161] Sobre el límite newtoniano puede consultarse, por ejemplo, los libros ya citados de A. Einstein "*El significado de la relatividad*", 2ª edición, Espasa Calpe Argentina, Buenos Aires, 1952; L. Landau y E. Lifchitz "*Teoría del campo*", Reverté, Barcelona, 1982; R. Lambourne, "*Relativity, gravitation and cosmology*", Cambridge University Press, U.K. 2010; W. Rindler "*Relativity: special, general and cosmological*", Oxford University Press, U.K. 2001; M. Hobson, G. Efstathiou y A. Lasenby, "*General relativity: an introduction for physicists*", Cambridge University Press, Cambridge, U.K. 2006; C. Misner, K. Thorne y J. Wheeler "Gravitation", W. Freeman, San Francisco, 1973.

[162] La partición del tensor métrico se efectúa de tal forma que las perturbaciones $h_{i,k}\left(x^0, x^1, x^2, x^3\right)$ no sean constantes en toda la variedad, a menos que se anulen en todas partes.

en forma significativa mediante colisiones.
Un sistema de cuerpos materiales lo bastante separados como para ser considerado en gran escala espacial como un fluido con presión nula en todas partes, forma lo que se conoce como **polvo**.
En el caso de un polvo las componentes del tensor de energía-impulsión se reducen a la forma $T_{i,k} = \rho \cdot U_i \cdot U_k$ siendo ρ la densidad del polvo y U_j las componentes del cuadrivector velocidad introducido en la parte 8.5: $U_0 = \gamma \cdot c$ $U_1 = \gamma \cdot v_1$ $U_2 = \gamma \cdot v_2$ $U_3 = \gamma \cdot v_3$ Como en este caso todas las componentes v_j de la velocidad local del polvo son por hipótesis pequeñas respecto de c, resulta que $\gamma \cong 1$ y que $T_{0,0} = \rho \cdot c^2$ es dominante respecto de las demás componentes del tensor, $T_{0,k}$ (con $k = 1,2,3$) y con mayor razón $T_{i,k}$ (con $i,k = 1,2,3$), que se pueden despreciar.
Entonces el escalar T , que aparece en las ecuaciones del campo gravitatorio puestas en la forma $R_{i,k} = -\kappa \cdot \left(T_{i,k} - \frac{1}{2} \cdot g_{i,k} \cdot T \right)$, verifica: $T \cong T_{0,0} = \rho \cdot c^2$
Como consecuencia, estas 10 ecuaciones se reducen a 9 ecuaciones homogéneas $R_{i,k} \cong 0$ cuando al menos uno de los subíndices no es nulo, y una única ecuación no homogénea, donde la densidad de energía asociada al polvo aparece como fuente del campo:

$$R_{0,0} = -\kappa \cdot \left(T_{0,0} - \frac{1}{2} \cdot g_{0,0} \cdot T \right) = -\frac{1}{2} \kappa \cdot \rho \cdot c^2$$

Es posible demostrar que, cuando las perturbaciones en el tensor métrico de la variedad $h_{i,k}$ respecto del tensor métrico del tiempo-espacio de Minkowski son débiles, entonces $R_{0,0} \cong -\sum_{i=0}^{3} \frac{\partial \Gamma^i{}_{0,0}}{\partial x^i}$ y $\Gamma^i{}_{0,0} \cong -\frac{1}{2} \sum_{j=0}^{3} m^{i,j} \cdot \frac{\partial h_{0,0}}{\partial x^j}$. Como consecuencia de estas dos últimas relaciones se obtiene: $R_{0,0} \cong \frac{1}{2} \sum_{i,j=0}^{3} m^{i,j} \cdot \frac{\partial^2 h_{0,0}}{\partial (x^i)^2}$
Pero las coordenadas locales se pueden seleccionar de modo que $m_{0,0} = +1$, $m_{1,1} = m_{2,2} = m_{3,3} = -1$ y $m_{i,k} = 0$ siempre que $i \neq k$.
Si, además, suponemos que las perturbaciones $h_{i,k}$ se mantienen aproximadamente constantes en el tiempo, resulta que las derivadas respecto de x_0 se pueden despreciar.
De donde, empleando coordenadas cartesianas, finalmente resulta que:

$$R_{0,0} \cong -\frac{1}{2} \left(\frac{\partial^2 h_{0,0}}{\partial (x)^2} + \frac{\partial^2 h_{0,0}}{\partial (y)^2} + \frac{\partial^2 h_{0,0}}{\partial (z)^2} \right)$$

Como $R_{0,0} \cong -\frac{1}{2} \kappa \cdot \rho \cdot c^2$ se obtiene para la componente $h_{0,0}$ una ecuación de Poisson:

$$\frac{\partial^2 h_{0,0}}{\partial (x)^2} + \frac{\partial^2 h_{0,0}}{\partial (y)^2} + \frac{\partial^2 h_{0,0}}{\partial (z)^2} \cong \kappa \cdot \rho \cdot c^2$$

Un polvo formado por cuerpos materiales con masa de reposo no nula origina líneas de universo temporaloides caracterizadas por tener un elemento de intervalo positivo.
En ese caso las ecuaciones de las líneas geodésicas que describen el movimiento de una

partícula de polvo debido solamente a efectos gravitatorios se pueden poner en términos del intervalo como parámetro: $\frac{d^2x^i(s)}{ds^2}+\sum_{j,k}\Gamma_{j,k}{}^{i}\cdot\frac{dx^j(s)}{ds}\cdot\frac{dx^k(s)}{ds}=0$

En condiciones de campo gravitatorio débil y velocidades de movimiento pequeñas, el examen de estas ecuaciones teniendo en cuenta que el intervalo se puede aproximar por la expresión $ds^2\cong\left(dx^0\right)^2-\sum_{j=1}^{3}(dx)^2=c^2\cdot dt^2-dx^2-dy^2-dz^2$, arroja las siguientes relaciones: $\frac{d}{ds}x^0\approx 1$ $\frac{d}{ds}x^i\approx\frac{v^i}{c}$ $\frac{d^2}{ds^2}x^i\approx\frac{1}{c^2}\cdot\frac{d^2}{dt^2}x^i\approx-\Gamma^i{}_{0,0}$ $(i=1,2,3)$

Como ahora $\Gamma^i{}_{0,0}\cong-\frac{1}{2}\sum_{j=0}^{3}m^{i,j}\cdot\frac{\partial h_{0,0}}{\partial x^j}=\frac{1}{2}\frac{\partial h_{0,0}}{\partial x^i}$ ($i=1,2,3$ con $m^{i,i}=-1$), entonces resulta la ecuación del movimiento $\frac{d^2}{dt^2}x^i\approx-\frac{c^2}{2}\frac{\partial h_{0,0}}{\partial x^i}$ que es idéntica a la fórmula newtoniana $\frac{d^2x^i}{dt^2}=-\frac{\partial\varphi}{\partial x^i}$ si igualamos $h_{0,0}=\frac{2}{c^2}\cdot\varphi$ siendo φ el potencial gravitatorio newtoniano.

Sustituyendo este resultado en $\frac{\partial^2 h_{0,0}}{\partial(x)^2}+\frac{\partial^2 h_{0,0}}{\partial(y)^2}+\frac{\partial^2 h_{0,0}}{\partial(z)^2}\cong\kappa\cdot\rho\cdot c^2$ se obtiene

$$\frac{\partial^2\varphi}{\partial(x)^2}+\frac{\partial^2\varphi}{\partial(y)^2}+\frac{\partial^2\varphi}{\partial(z)^2}=\frac{c^4}{2}\cdot\rho\cdot\kappa$$

De la comparación con la ecuación del potencial gravitatorio newtoniano $\frac{\partial^2\varphi}{\partial(x)^2}+\frac{\partial^2\varphi}{\partial(y)^2}+\frac{\partial^2\varphi}{\partial(z)^2}=4\pi\cdot G\cdot\rho$ resulta la fórmula, introducida en 10.2.3, que conecta la constante κ de la ecuación tensorial para el campo gravitatorio con la constante de gravitación de Newton G y con la velocidad de la luz en el vacío:

$$\kappa=\frac{8\pi\cdot G}{c^4}$$

Admitiendo esta última relación, y la identificación $h_{0,0}=\frac{2}{c^2}\cdot\varphi$, se obtiene la teoría newtoniana de la gravitación como caso límite de la teoría relativista para campos débiles y velocidades pequeñas: es decir, la ecuación newtoniana para el potencial del campo gravitatorio y las ecuaciones newtonianas del movimiento de los cuerpos en ese campo.

En el límite newtoniano las componentes del campo gravitatorio son:

$-\frac{\partial\varphi}{\partial x^i}=-c^2\cdot\Gamma^i{}_{0,0}$ $(i=1,2,3)$ En este caso particular los símbolos $\Gamma^i{}_{0,0}$ se pueden considerar como representación de las componentes del campo gravitatorio.

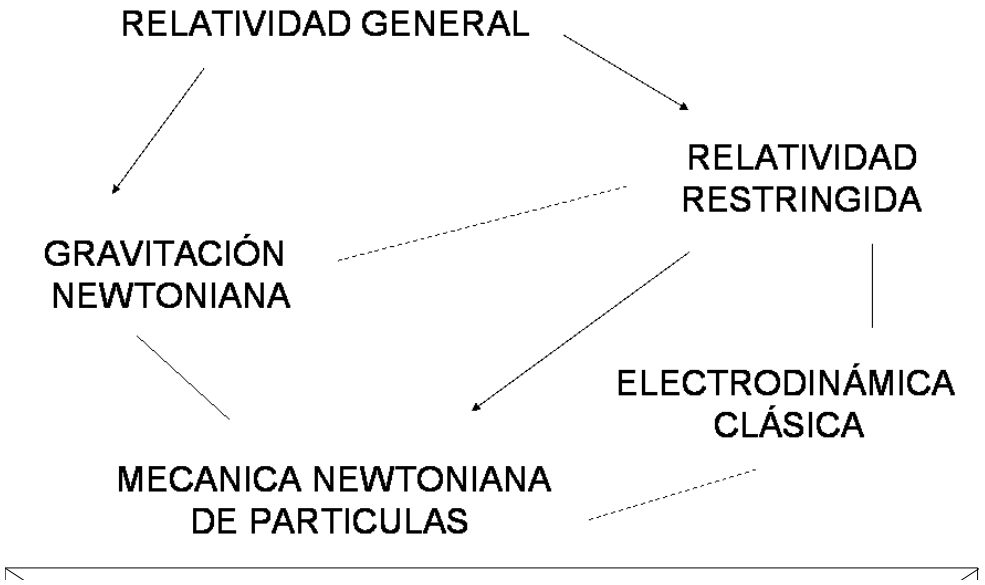
En la aproximación que conduce a la gravitación newtoniana $g_{00}=1+h_{0,0}=1+\frac{2}{c^2}\cdot\varphi$, mientras que $g_{11}=g_{22}=g_{33}=-1$ y $g_{ik}=0$ cuando $i\neq k$. Entonces el elemento de

intervalo se puede rescribir así: $ds^2 = \left(1 + \frac{2 \cdot \varphi(t,x,y,z)}{c^2}\right) \cdot c^2 \cdot dt^2 - \left(dx^2 + dy^2 + dz^2\right)$

Este resultado se interpreta a veces diciendo que en la aproximación newtoniana, deducida en el marco de la teoría generalizada, el tiempo está curvado.
En todo caso, para un elemento de intervalo como éste, en el que el potencial gravitatorio varía con las coordenadas espaciales, la correspondiente variedad tetra-dimensional de tiempo-espacio-materia no es plana. Siempre que los efectos gravitatorios no puedan ser despreciados, el tiempo-espacio estará curvado: así el tiempo-espacio-materia que constituye el caso límite consistente con la teoría newtoniana de la gravitación es una variedad curvada, no es plana, pese a que las velocidades de los cuerpos que forman los sistemas gravitantes son pequeñas respecto de la velocidad de la luz.
Otra consecuencia notable de este caso límite es el efecto del campo gravitatorio sobre el tiempo propio que se considera en el próximo capítulo, en la parte 11.3.

10.4 Las relaciones epistemológicas entre las teorías de la física clásica luego de ser aceptada la teoría generalizada de la relatividad

A partir de 1916 tenemos, en lo que puede denominarse teoría clásica (o sea, no cuántica) de campos y partículas, un panorama como el representado en el esquema siguiente:



Como en los esquemas previos, las flechas de trazo continuo representan relaciones de dominio de una teoría sobre otra. La teoría generalizada domina sobre la gravitación newtoniana y sobre la teoría restringida. Las demás relaciones de dominio, compatibilidad e inconsistencia ya fueron explicadas.

Como vimos en 10.3, si la curvatura del continuo tiempo-espacio-materia tiende a cero en todas partes, ese continuo se aproxima al tiempo-espacio de Minkowski de la relatividad restringida, el campo gravitatorio desaparece y la formulación matemática de las leyes físicas no vinculadas con la gravitación se transforma en su formulación acorde con la relatividad restringida. Cuando la curvatura del continuo tiende a cero, la

teoría generalizada tiene como caso límite a la teoría restringida.

El movimiento relativista de cuerpos materiales en campos gravitatorios débiles y con velocidades pequeñas en comparación con la velocidad de la luz se aproxima al movimiento según la mecánica clásica en un potencial newtoniano: la teoría de la gravitación de Newton se obtiene como caso límite de la teoría de la gravitación de –Einstein. El tiempo-espacio-materia correspondiente a este caso límite está (débilmente) curvado, y solo se aplana si desaparecen por completo los efectos de la gravedad. Esto es una consecuencia de la implementación matemática de la idea del movimiento bajo los efectos de un campo gravitatorio considerado como movimiento libre en un espacio curvo, equiparando gravitación con curvatura.
En consecuencia, al igual que en el caso de las relaciones epistemológicas entre la mecánica newtoniana y la mecánica relativista consideradas en 8.8, también en el caso de la teoría newtoniana de la gravitación y la teoría relativista de la gravitación no hay una traducibilidad estricta entre los lenguajes de ambas teorías. El entrelazamiento de los conceptos de tiempo y espacio ya presente en la teoría restringida, más la novedad de la curvatura, propia de la teoría generalizada y que permanece en el caso límite correspondiente a la gravitación newtoniana, muestran que consideradas en sentido estricto esas dos teorías son inconmensurables. No obstante hay suficientes homologías entre los términos propios de cada uno de los lenguajes como para poder realizar una comparación entre teorías sobre la base de sus balances predictivos y concluir que la nueva teoría relativista de la gravitación remplaza con propiedad a la antigua teoría.

Puesto que según la teoría generalizada de la relatividad, un campo gravitatorio representa un cambio en las propiedades geométricas del continuo tiempo-espacio, cabe esperar que los intervalos de tiempo de los procesos y las distancias entre los cuerpos se vean afectados por el campo. Esto se confirma al analizar la relación entre el elemento de línea y los cambios en las coordenadas de tiempo y espacio cuando varían las componentes del tensor métrico asociadas con una variación del campo gravitatorio. Como cabe esperar que este último se modifique a su vez no solo de un lugar a otro sino de un instante a otro, los cuerpos materiales que componen un sistema mecánico en principio no permanecen a distancias fijas unos de otros: la deformación es inevitable. Los conceptos de equidistancia o de velocidad relativa fija entre dos cuerpos, en general carecen de sentido en presencia de un campo gravitatorio dinámico. Entonces los sistemas para la ubicación espacial de eventos, construidos mediante cuerpos cuyas distancias mutuas pueden ser tan grandes como se quiera y se mantienen invariables, que se postulan como parte de los marcos inerciales de la relatividad restringida, ya no son concebibles en la teoría generalizada.
Para situar eventos en el tiempo-espacio en presencia de un campo gravitatorio se precisarían infinitos cuerpos con sus correspondientes relojes marchando con sus propios ritmos y llenando todo el continuo: solamente de esta forma se tendría un marco de referencia adecuado a la teoría generalizada de la relatividad.
No obstante, en presencia de un campo gravitatorio, pero **localmente** en el tiempo-espacio, los marcos de referencia de la relatividad restringida se pueden utilizar como una buena aproximación, incluyendo los relojes (que denominamos newtonianos en 2.1) y las varillas supuestamente rígidas, que se utilizan para determinar relaciones métricas

utilizando un modelo de espacio euclidiano[163].
También en forma local se verifica con muy buena aproximación que el movimiento de una partícula libre de fuerzas respecto de uno de estos marcos es rectilíneo y uniforme en el espacio euclidiano correspondiente. Pero si se abandona la escala local y se mantiene la descripción en un espacio euclidiano, el movimiento no puede ya considerarse con buena aproximación ni rectilíneo ni uniforme, debido al efecto del campo gravitatorio.

10.5 El significado de la teoría matemática de la relatividad

Para finalizar este capítulo sobre la formulación matemática de la teoría relativista de la gravitación, resumiremos algunas de sus aspectos de mayor importancia.
El nuevo continuo tiempo-espacio introducido en la teoría generalizada, se puede representar localmente por un continuo de Minkowski, de forma análoga a como una superficie curva lo bastante regular contenida el espacio euclidiano tridimensional se puede aproximar en cada uno de sus puntos por el correspondiente plano tangente.
La curvatura de este espacio en cada punto depende de la distribución de masas y sus movimientos, de tal forma que sus propiedades geométricas no pueden separarse de la materia ponderable y su movimiento. Pero además, el campo gravitatorio depende del campo electromagnético y en general de cualquier sistema físico: puede decirse que la gravitación emana de todas las entidades físicas[164].
En principio, si se conocen y se agregan las ecuaciones de movimiento no gravitatorio de la materia y la radiación a la ecuación tensorial de la gravitación, esta última se acoplaría con las primeras a través de la versión generalizada del tensor de energía-impulsión.
Pero, a su vez, las modificaciones en las propiedades métricas del tiempo-espacio provocadas por el movimiento de la materia ponderable y la radiación (a través del tensor generalizado de energía-impulsión que aparece en la ecuación tensorial del campo gravitatorio) actúan sobre ese movimiento: se puede concebir así el cierre de una cadena causal de tal forma que la dinámica de la materia-energía y la geometría del tiempo-espacio curvo se determinen recíprocamente. Se tiene en última instancia un mundo físico visto como una totalidad tiempo-espacio-energía o en forma equivalente, una totalidad de tiempo-espacio-materia.
Al final 10.2.4 se llamó la atención sobre el siguiente hecho: de las ecuaciones del campo gravitatorio, luego de varias transformaciones se pueden deducir las ecuaciones del movimiento de la materia y la radiación que produce el campo (las ecuaciones de las geodésicas en el tiempo-espacio-materia). Esto constituye una diferencia notable con la teoría electromagnética, en la cual las ecuaciones de la teoría implican la ley de conservación de la carga pero no las ecuaciones del movimiento de las cargas que producen el campo. Mientras que la ubicación y el estado de movimiento de las cargas eléctricas puede ser dado en forma independiente del campo electromagnético que esas cargas producen (por lo menos en principio), en el caso del campo gravitatorio la ubicación y el estado de movimiento de las fuentes del campo (partículas y radiación) no se pueden establecer en forma arbitraria: deben ser determinados al mismo tiempo

[163] El campo gravitatorio local se sustituye por un marco de referencia no inercial localmente equivalente, como se explicó en 10.1, y el marco no inercial es aproximado a su vez en el entorno de un instante de tiempo por un marco inercial.

[164] Tanto las partículas como los campos electromagnéticos, incluyendo las ondas electromagnéticas, y los mismos campos gravitatorios y las ondas de gravitación producen campos gravitatorios.

que el campo. Como las ecuaciones de la gravitación solo determinan el movimiento debido a la gravedad, en general es necesario introducir algunas relaciones adicionales específicas para resolver el problema. Por ejemplo, si las componentes del tensor de energía-impulsión corresponden a un fluido ideal, se debe introducir una relación entre la presión y la densidad de energía de ese fluido, como se verá en el capítulo 12 a propósito de los modelos cosmológicos.
Otra peculiaridad de la teoría generalizada de la relatividad es el rol del campo gravitatorio en las leyes de conservación. Como vimos en 9.5.2, las derivadas covariantes involucran los coeficientes de conexión $\Gamma^{j}{}_{k,l}$ de la variedad. Para el tensor de energía-impulsión: $D_k T^{i,j} = \frac{\partial}{x^k} T^{i,j} + \sum_{l=1}^{n} \Gamma^{j}{}_{k,l} \cdot T^{l,j} + \sum_{l=1}^{n} \Gamma^{j}{}_{k,l} \cdot T^{i,l}$ $(i, j, k = 0,1,2,3)$
Como vimos en 10.2.3, los coeficientes de conexión se pueden considerar como las componentes del campo gravitatorio. Resulta entonces que en la expresión matemática de la ley de conservación de la energía-cantidad de movimiento $\sum_{l=0}^{3} D_l T^{i,l} = 0$ aparecen las componentes del campo gravitatorio. No puede formularse un principio de conservación para la materia aislada del campo gravitatorio. El campo gravitatorio transfiere energía y cantidad de movimiento lineal y angular a la materia, y a su vez debe recibir de esta última energía y cantidad de movimiento lineal y angular[165].

Las ecuaciones del campo gravitatorio no son lineales. Como consecuencia el principio de superposición, que es aplicable al campo electromagnético en el vacío, y que permite obtener nuevas soluciones sumando soluciones conocidas multiplicadas por coeficientes numéricos arbitrarios, ya no es aplicable en la teoría relativista de la gravitación.
Esta no linealidad dificulta o imposibilita la construcción de soluciones analíticas para resolver problemas concretos[166] y conduce a la necesidad de emplear técnicas de simulación digital de las soluciones mediante ordenadores. Estas últimas, dada la naturaleza de las ecuaciones, no están exentas de dificultades, tanto en relación con los algoritmos numéricos utilizados como en relación con su implementación efectiva en el ordenador. Se han hallado varias dinámicas caóticas en la teoría relativista de la gravitación, con las consecuencias (mencionadas en la parte 3.4) sobre la imposibilidad de predicción, en la práctica, de los detalles cuantitativos de esas dinámicas.

La formulación matemática de la teoría relativista de la gravitación es local y no global: determina las características locales del continuo tiempo-espacio-materia mediante un sistema de ecuaciones diferenciales en derivadas parciales que por sí solas no determinan la estructura en gran escala del espacio y el tiempo, es decir, sus

---

[165] A. Einstein "*El significado de la relatividad*", Espasa Calpe Argentina, Buenos Aires, 1952.

[166] De todos modos, se han obtenido muchas soluciones analíticas tentativas sustituyendo campos tensoriales, que se consideraban como posibles candidatos a métrica para el continuo tiempo-espacio, en el primer miembro y analizando si el segundo miembro que se obtenía podía o no ser un tensor de energía-impulsión razonable. Los dos criterios más utilizados para clasificar soluciones son la clasificación algebraica de Petrov basada en un tensor introducido por Weyl y la clasificación según los grupos de transformaciones que conservan la estructura geométrica del espacio, basada en los denominados campos de Killing. Una explicación breve de estas dos modalidades de clasificación de los campos gravitatorios puede hallarse en D. Ivanenko y G. Sardanashvili "Gravitación", capítulo 1, Editorial URSS, Moscú, 2005.

propiedades globales.
Como consecuencia las ecuaciones del campo gravitatorio poseen soluciones muy variadas, que corresponden a continuos tiempo-espacio-materia con propiedades globales muy diferentes. En la jerga cosmológica se hace referencia a **soluciones con topologías diferentes.**
Por ejemplo, cuando las curvaturas locales son positivas, el continuo tiempo-espacio-materia puede resultar finito aun cuando sea ilimitado. Y no solamente finito en relación con las coordenadas espaciales, sino finito también en relación con la coordenada temporal. En este espacio curvado la asimetría entre el pasado y el futuro se verifica localmente, pero puede no verificarse globalmente, ya que una línea de universo que representa una trayectoria posible de una partícula, puede cerrarse sobre sí misma, como lo demostró Kurt Gödel en un famoso artículo sobre las soluciones de las ecuaciones de campo en la teoría general de la relatividad.
La única forma de excluir esta posibilidad es añadiendo a la teoría un principio ad-hoc que la elimine, como sugirió Einstein mismo, o bien modificar desde su base misma la teoría generalizada.

11. Solución de Schwarzschild y confirmaciones de la teoría de la gravitación

"Cuando nos encontramos con un sistema de coordenadas que incluye una coordenada t y una coordenada r, resulta tentador asumir que t debe representar el tiempo y r la distancia radial desde el origen. No obstante, una suposición como ésta es siempre peligrosa y a menudo es errónea.
El hecho simple es que en relatividad generalizada las coordenadas son sistemas de marcas arbitrarias elegidas para distinguir un evento de otro. Esto nos otorga una gran libertad sobre cómo definir las coordenadas, una libertad que explotamos en la deducción de la métrica de Schwarzschild. La relación entre las diferencias de coordenadas que separan eventos y los correspondientes intervalos de tiempo o de distancia que serían medidos por un observador específico debe ser trabajada utilizando la métrica del tiempo-espacio. No se puede suponer que los tiempos y distancias "físicos" que se medirían utilizando relojes y varillas, se encuentran directamente especificados por las coordenadas. ***En relatividad general, las coordenadas no poseen significado métrico inmediato****.*
Einstein consideraba que esta característica de la relatividad generalizada era un motivo de perplejidad. En su propio testimonio sobre cómo fue el desarrollo de la teoría a partir de 1908, decía: "¿Por qué se precisaron siete años para construir la teoría generalizada de la relatividad? ***La razón principal yace en que no es fácil liberarse de la idea de que las coordenadas deben tener un significado métrico inmediato****". (Citado en P.Schlipp (Ed.) Albert Einstein-Philosopher Scientist, 3ª Edición, Open Court, Illinois, 1969)."*
R. Lambourne en "Relativity, gravitation and cosmology", p. 115, Cambridge University Press, Cambridge, U.K., 2010.

La obtención de soluciones analíticas de las 10 ecuaciones en derivadas parciales no lineales y acopladas que describen el campo gravitatorio, con la condición adicional de que sean de interés desde el punto de vista físico para resolver un problema planteado de antemano, es como dijimos en 10.5 una tarea difícil, y a veces imposible, aun cuando se suponga conocido el tensor de energía impulsión. Por otra parte, una vez descubierta una candidata a solución, es preciso asegurarse que no sea una solución ya hallada expresada en otro sistema de coordenadas.

En 1916, poco después de publicado el trabajo de Einstein en el que aparecían las ecuaciones del campo gravitatorio, Karl Schwarzschild, por ese entonces participando en la Primera Guerra Mundial y destinado al frente de batalla en calidad de oficial de artillería, se hizo de tiempo para hallar la primer solución exacta de esas ecuaciones, y unos meses más tarde falleció. En forma independiente, Johannes Droste halló la misma solución y la presentó, también en 2016, en su tesis de doctorado[167].

[167] Karl Schwarzschild, On the Gravitational Field of a Mass Point according to Einstein's Theory, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin, Phys.-Math. Klasse 1916, 189-196. (Traducción al inglés disponible en archive.org) y Johannes Droste, The Field of a Single Centre in Einstein's Theory of Gravitation, and the Motion of a Particle in that Field, Koninklijke Nederlandsche Akademie van Wetenschappen Proceedings 19, 197 (1917).

Con el tiempo, la versión de la solución de Schwarzschild debida a David Hilbert[168] resultó ser la más conocida y aplicada de todas las soluciones analíticas conocidas para el campo gravitatorio.
Ha servido como base matemática para calcular efectos que se han utilizado para confirmar la teoría de la gravitación de Einstein: la precesión del perihelio de Mercurio, la deflexión de la luz por el Sol, el corrimiento al rojo y la dilatación del tiempo debido al campo gravitatorio, y el retardo de los pulsos de radar en el sistema solar.
También se la utiliza para calcular algunos de los efectos asociados a las denominadas lentes gravitacionales y varias propiedades que se atribuyen a los agujeros negros. A partir de 1963, fue extendida en sucesivas etapas por otros investigadores, en buena medida interesados en la física de los agujeros negros.
En este capítulo se estudian algunas propiedades y algunas aplicaciones de la mencionada versión de Hilbert de la solución de Schwarzschild. En el capítulo 13 se considera en forma somera su aplicación a los agujeros negros.
Comenzaremos con un modelo que permite hallar, en forma directa y simple, las componentes del tensor métrico asociado a una masa puntual en el vacío. A partir de este resultado aproximado, mediante una nueva aproximación que compensa el error de cálculo de la primera, deduciremos la métrica de Schwarzschild.

11.1 El cálculo del tensor métrico mediante el modelo de Lenz. Métrica de Schwarzschild. Singularidades de coordenadas y singularidades gravitacionales.

Es posible hallar las componentes $g_{i,k}$ del tensor métrico en algunos casos particulares sin necesidad de resolver las ecuaciones de la gravitación, utilizando solamente una combinación de relatividad restringida con el principio de equivalencia entre la masa inercial y la masa gravitatoria, en forma análoga a la ya empleada en 10.1 con el fin de analizar el sincronismo de relojes, en reposo relativo, pero ubicados en puntos distintos de un campo gravitatorio que se supone se puede describir con un único potencial escalar, calculable en forma aproximada mediante la teoría newtoniana.
Un modelo muy simple para hallar el tensor métrico utilizando estas hipótesis simplificadoras es el de W. Lenz, comunicado por este autor al Profesor Sommerfeld en 1944[169].
Lenz comenzaba por considerar un campo gravitatorio simétrico producido por una masa central M que se consideraba como una masa puntual, en reposo respecto de un marco de referencia K, sujeto al campo de gravitación
Para describir los eventos respecto de este marco K usaba un sistema de coordenadas esféricas $x^1 = r, x^2 = \theta, x^3 = \phi$ con centro en la masa puntual M, y el instante de

[168] D. Hilbert, Gravitationsfield eines Massenpunktes, Nachr. Ges. Wiss. Göttingen, Math. Phys. Kl, vol.53, 1917.

[169] A. Sommerfeld "*Electrodynamics*", Academic Press, New York, 1952, pp. 313-315. En el libro de J. Palacios "*Relatividad*", Espasa Calpe, Madrid, 1960, se puede encontrar una crítica y una corrección a la presentación que hace Sommerfeld del modelo de Lenz.

tiempo $x_0 = t$ determinado utilizando el correspondiente reloj sincronizado en K, de modo que cada evento se caracteriza por la cuaterna $(x_0, x_1, x_2, x_3) = (ct, r, \theta, \phi)$

Luego Lenz considera una caja solidaria a un marco K∞ que cae libremente en dirección a M. Como K∞ se encuentra en caída libre en el campo gravitatorio producido por M, no presenta campo gravitatorio alguno y transporta consigo una métrica de Minkowski $g_{00} = 1, g_{11} = -1, g_{22} = -1, g_{33} = -1, g_{i,k} = 0 \; i \neq k$, válida también en el marco K cuando la distancia radial a la masa tiende a infinito y el campo gravitatorio se desvanece. La coordenada x^∞ se dirige radialmente hacia la masa puntual mientras que las coordenadas y^∞ y z^∞ se disponen transversales a esa dirección. Por su parte t^∞ representa el instante determinado por el correspondiente reloj sincronizado en K∞.

Supongamos que v es la magnitud de la velocidad de caída de K∞ hacia M cuando la distancia radial entre la caja y la masa es r. Tanto r como v se toman respecto del marco K, sujeto al campo gravitatorio de M.

Lenz asume que K∞ (que no presenta campo gravitatorio) se puede considerar instantáneamente como si fuera un marco de referencia inercial que se mueve con velocidad v respecto del marco K en reposo. En ese caso se aplican las transformaciones de Lorentz escritas de la siguiente forma infinitesimal[170]:

$$dx^\infty = \frac{dr}{\sqrt{1-\frac{v^2}{c^2}}} \qquad dt^\infty = \sqrt{1-\frac{v^2}{c^2}} \cdot dt \qquad dy^\infty = r \cdot d\theta \qquad dz_\infty = r \cdot sen\theta \cdot d\phi$$

Ahora escribamos el intervalo respecto del marco K∞, es decir:

$$ds^2 = c^2 \cdot dt^{\infty 2} - dx^{\infty 2} - dy^{\infty 2} - dz^{\infty 2}$$

Respecto de K adopta la forma:

$$ds^2 = \left(1-\frac{v^2}{c^2}\right) \cdot c^2 \cdot dt^2 - \frac{dr^2}{\left(1-\frac{v^2}{c^2}\right)} - r^2 \cdot d\theta^2 - r^2 \cdot sen^2\theta \cdot d\phi^2$$

Supongamos ahora que la caja que cae hacia la masa gravitante M posee una pequeña masa m cuando se halla en movimiento respecto de K∞. Designemos con m_0 a su masa en reposo y supongamos que cuando la distancia entre las dos masas tiende a infinito, la velocidad de la caja respecto del marco K tiende a cero. En ese caso, respecto de K, la energía de la caja a distancia infinita de M es su energía de reposo $m_0 \cdot c^2$.

---

[170] La contracción del elemento radial y la dilatación del tiempo que emplea Lenz se desprenden de los resultados obtenidos en 6.1. El elemento dx^∞ se encuentra en reposo en K∞ por lo cual su longitud dr respecto de K debe medirse simultáneamente en K: $dr = \left(\sqrt{1-\frac{v^2}{c^2}}\right) \cdot dx^\infty$ El elemento dt^∞ se mide respecto a un reloj fijo a K∞ por lo cual es el tiempo propio de K∞ que se vincula con el elemento de tiempo dt medido en K mediante $dt = \frac{dt^\infty}{\sqrt{1-\frac{v^2}{c^2}}}$.

Lenz supuso luego que se verifica la siguiente forma de la ecuación de la energía respecto del sistema K: $(m-m_0)\cdot c^2 - G\cdot\frac{m\cdot M}{r}=0$

La energía cinética de la caja viene dada por $(m-m_0)\cdot c^2$ mientras que su energía potencial gravitatoria viene dada en una primera aproximación por la expresión newtoniana $-G\cdot\frac{m\cdot M}{r}$.

Como a distancia infinita de la masa gravitante M se verifica $m=m_0$, si se asume que la energía potencial gravitatoria se anula en el infinito, se ve que la suma de la energía cinética y potencial gravitatoria de la caja se debe igualar a cero.

Teniendo en cuenta que $m=\frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$, de la ecuación de la energía se desprende que $\sqrt{1-\frac{v^2}{c^2}}=1-\frac{\alpha}{r}$ donde $\alpha=\frac{G\cdot M}{c^2}$. Si M es igual a la masa del sol, entonces α es de aproximadamente 3 km. En consecuencia, como cabe esperar que para el caso del sol el modelo se aplique a distancias varios órdenes mayores que el radio solar (distancias del orden de la distancia de la Tierra al Sol) se puede efectuar la muy buena aproximación:

$$\left(1-\frac{v^2}{c^2}\right)=\left(1-\frac{\alpha}{r}\right)^2\approx 1-\frac{2\cdot\alpha}{r}$$

A partir de esa aproximación se obtiene para el intervalo:

$$ds^2=\left(1-\frac{2\cdot\alpha}{r}\right)\cdot c^2\cdot dt^2-\frac{dr^2}{\left(1-\frac{2\cdot\alpha}{r}\right)}-r^2\cdot d\theta^2-r^2\cdot sen^2\theta\cdot d\phi^2$$

Ésta es la famosa expresión, que puede deducirse [171] a partir de construir una solución de las ecuaciones de la gravitación de Einstein $R_{ik}-\frac{1}{2}\cdot g_{ik}\cdot R=0$ para la región vacía del tiempo-espacio que rodea a un cuerpo puntual[172] de masa M, en reposo y sin rotación respecto de un marco K.

Para esa métrica:

$$g_{00}=1-\frac{2\cdot\alpha}{r}\qquad g_{11}=-\frac{1}{1-\frac{2\cdot\alpha}{r}}\qquad g_{22}=-r\qquad g_{33}=-r\cdot sen\theta\qquad g_{i,k}=0\quad i\neq k$$

Es una métrica **estacionaria**, puesto que las componentes del tensor métrico no dependen de la coordenada temporal. Además es **estática**: permanece invariante bajo la

[171] Una deducción bastante detallada de la versión de Hilbert de la métrica de Schwarzschild puede hallarse en R. Lambourne "*Relativity, gravitation and cosmology*", capítulo 5, Cambridge University Press, Cambridge, U.K., 2010. A nivel más avanzado se puede encontrar en el capítulo 25 de C. Misner, K. Thorne y J. Wheeler "*Gravitation*" W. Freeman, San Francisco, 1973, y en el párrafo 97 de L. Landau y E. Lifchitz, "*Teoría del campo*", Reverté, Barcelona, 1982.

[172] O al menos con una masa distribuida con simetría esférica.

inversión del tiempo $t \to -t$, puesto que los $g_{i,k}$ no dependen del tiempo y en la métrica no aparecen términos mixtos[173] del tipo $g_{t,r} \cdot dt \cdot dr \quad g_{t,\theta} \cdot dt \cdot d\theta \quad g_{t,\phi} \cdot dt \cdot d\phi$.
El tiempo-espacio correspondiente a esta métrica es **asintóticamente plano**:
Cuando $r \to \infty$, se ve que $g_{00} \to 1, g_{11} \to -1$ mientras que $g_{22} = -r, g_{33} = -r \cdot sen\theta$, y $g_{i,k} = 0$ para $i \neq k$: al alejarnos de la masa gravitante recuperamos en K la métrica de Minkowski propia de un marco inercial: $ds^2 = c^2 \cdot dt^2 - dr^2 - r^2 \cdot d\theta^2 - r^2 \cdot sen^2\theta \cdot d\phi^2$
Además ese tiempo-espacio posee **simetría esférica** en el siguiente sentido:
Si fijamos un instante $t = T$ y un valor $r = R$, el intervalo se reduce a $ds^2 = -R^2 \cdot d\theta^2 - R^2 \cdot sen^2\theta \cdot d\phi^2$. Pero en este caso particular $-ds^2$ describe la geometría sobre la superficie de una esfera bidimensional inmersa en un tiempo-espacio tetra-dimensional, de radio R y centrada en la masa generadora del campo. Desde el punto de vista físico y con esa métrica, todos los puntos del tiempo-espacio $(c \cdot T, R, \theta, \phi)$ correspondientes a esa esfera son equivalentes.

El denominado radio de Schwarzschild $r_S = 2 \cdot \alpha = 2 \cdot \dfrac{G \cdot M}{c^2}$ establece un límite inferior a las distancias radiales para las cuales se puede aplicar la fórmula para el intervalo en las coordenadas originales. Una vez alcanzado este límite se anula el coeficiente g_{00} y se hace infinito el coeficiente g_{11} del tensor métrico. No obstante no se trata de una singularidad físicamente significativa, o sea, gravitacional, sino de una singularidad en la coordenada. A diferencia de lo que ocurre en el caso de una singularidad gravitacional, las **singularidades en las coordenadas** se pueden hacer desaparecer mediante un cambio de coordenadas adecuado[174].

La singularidad en $r = 0$ que presenta la métrica es una verdadera **singularidad gravitacional**, caracterizada porque al aproximarse a ella los invariantes que describen la curvatura del tiempo-espacio tienden a infinito. La singularidad no se puede eliminar mediante cambios de coordenadas[175]. Como se verá en el capítulo 13, lo que ocurre al interior del radio r_S corresponde a la física de los agujeros negros.

11.2 Gravitación y curvatura del continuo tiempo-espacio

La fórmula del intervalo se puede escribir en términos del radio de Schwarzschild:

---

[173] Estos términos mixtos aparecen en la métrica cuando la masa generadora del campo rota respecto del marco K. Si la rotación es uniforme, las componentes del tensor métrico no dependen del tiempo pero la métrica no es estática.

[174] Como el introducido por M. Kruskal en 1960, a partir de un análisis crítico de la asimetría en las así llamadas coordenadas avanzadas de Eddington- Finkelstein. Una explicación muy clara de estas coordenadas, en el contexto de un estudio sobre los agujeros negros, puede encontrarse en el capítulo 6 de R. Lambourne "*Relativity, gravitation and cosmology*", Cambridge University Press, Cambridge, U.K., 2010.

[175] C. Misner, K. Thorne y J. Wheeler "*Gravitation*" W. Freeman, San Francisco, 1973.

$$ds^2 = \left(1 - \frac{r_S}{r}\right) \cdot c^2 \cdot dt^2 - \frac{dr^2}{\left(1 - \frac{r_S}{r}\right)} - r^2 \cdot d\theta^2 - r^2 \cdot sen^2\theta \cdot d\phi^2$$

De la parte dependiente de las coordenadas espaciales $\frac{dr^2}{\left(1 - \frac{r_S}{r}\right)} + r^2 \cdot d\theta^2 + r^2 \cdot sen^2\theta \cdot d\phi^2$ se desprende que la traza en el continuo tiempo-espacio para un mismo instante de tiempo, es decir la parte del continuo que conocemos como el espacio físico, no es euclidiano.

Para ver esto con más detalle, supongamos que hacemos $d\phi = 0$. Entonces la parte espacial del intervalo es $\frac{dr^2}{\left(1 - \frac{r_S}{r}\right)} + r^2 \cdot d\theta^2$

Si calculamos la longitud de una línea equidistante de la masa central (circunferencia), es decir, para $dr = 0$, el elemento tangencial de línea se reduce a $r \cdot d\theta$. De aquí resulta que la longitud de la circunferencia es $2 \cdot \pi \cdot r$

Ahora hacemos $d\theta = 0$ y calculamos la distancia $d_{1,2}$ entre dos puntos de coordenadas r_1 y r_2, con $r_S < r_1 < r_2$: $d_{1,2} = \int_{r_1}^{r_2} \frac{dr}{\sqrt{\left(1 - \frac{r_S}{r}\right)}}$ Esta distancia es mayor que $r_2 - r_1$.

En la geometría euclidiana el cociente entre la diferencia de longitudes de dos circunferencias y la diferencia entre sus radios vale siempre $\frac{2\pi(r_2 - r_1)}{(r_2 - r_1)} = 2\pi$, mientras que por efecto del campo gravitatorio vale $\frac{2\pi(r_2 - r_1)}{d_{1,2}} < 2\pi$. Como $\frac{(r_2 - r_1)}{d_{1,2}}$ disminuye desde las proximidades de 1 cuando r_1 es muy grande respecto del radio de Schwarzschild hasta las proximidades de 0 cuando r_1 y r_2 se aproximan a ese radio crítico, el espacio físico, que es plano muy lejos de la masa gravitante, se curva cada vez más al aproximarnos a ella.

G. Birkhoff, en 1923, demostró que todas las soluciones con simetría esférica de las ecuaciones del campo gravitatorio en el vacío (es decir, en la región exterior a las fuentes del campo) y cuya métrica tiende en el infinito hacia una métrica plana, son equivalentes (a menos de un cambio de coordenadas) a la solución de Schwarzschild[176]. Este resultado se verifica aun cuando las fuentes del campo no se mantengan estacionarias, siempre que su efecto sea el mismo en todas direcciones. La solución de Scwarzschild, pese a que depende solamente de la masa total del cuerpo generador del campo, es de índole muy **general**.

Una consecuencia de este resultado de Birkhoff es que si una distribución de masa con simetría esférica se expande, se contrae o pulsa en forma estrictamente radial, en la

[176] G. Birkhoff"*Relativity and modern physics*", Harvard University Press, Cmbridge, Mass., 1923.

región exterior a cualquier esfera que contenga a la masa no se detectan modificaciones en el campo gravitatorio.

Como los radios R de las estrellas son de un orden de magnitud mucho mayor que el radio crítico r_S, se puede utilizar la solución de Schwarzschild en el espacio que rodea a una estrella, hasta la superficie de la misma.
En ese caso el cociente entre la longitud del ecuador y el radio de una estrella se puede estimar por la fórmula[177]:

$$\frac{longitud \cdot ecuador}{radio \cdot estelar} = \frac{2\pi \cdot R}{\int_{r_S}^{R} \frac{dr}{\sqrt{1-\frac{r_S}{r}}}} \approx 2\pi \cdot \left(1 - \frac{G \cdot M}{3 \cdot R \cdot c^2}\right)$$

Sobre la superficie del sol el apartamiento de 2π es de solamente 0,0000004.

Si ahora consideramos el flujo del tiempo en un mismo punto ($dr = 0, d\theta = 0, d\phi = 0$) tenemos que el intervalo se reduce a: $ds^2 = \left(1 - \frac{r_S}{r}\right) \cdot c^2 \cdot dt^2$

En términos de tiempo propio: $d\tau = \sqrt{\left(1 - \frac{r_S}{r}\right)} \cdot dt$

Entonces, si nos alejamos lo suficiente de la masa gravitante $d\tau \approx dt$, pero a medida que nos acercamos y el campo gravitatorio deja sentir sus efectos, dt se incrementa más respecto de $d\tau$: el campo gravitatorio produce una dilatación del tiempo.

Finalmente, consideremos un rayo de luz que se desplaza en las inmediaciones de la masa que genera el campo gravitatorio.
En ese caso el intervalo se anula siempre, debido a lo cual, respecto del marco K:

$$\left(1 - \frac{r_S}{r}\right) \cdot c^2 - \frac{\left(\frac{dr}{dt}\right)^2}{\left(1 - \frac{r_S}{r}\right)} - r^2 \cdot \left(\frac{d\theta}{dt}\right)^2 - r^2 \cdot sen^2\theta \cdot \left(\frac{d\phi}{dt}\right)^2 = 0$$

Tomando $d\theta = 0, d\phi = 0$, o sea, para una trayectoria radial de la luz, se obtiene una velocidad de propagación expresada en términos de la coordenada radial r:

$$\left|\frac{dr}{dt}\right| = \left(1 - \frac{r_S}{r}\right) \cdot c$$

Como la coordenada radial no representa una distancia (no posee un significado métrico inmediato), esta velocidad aparente no corresponde a la velocidad radial de la luz definida a través de la distancia recorrida por unidad de tiempo transcurrido.
Tomando $dr = 0, d\varphi = 0$, o sea, para una trayectoria en la que los fotones se mueven tangencialmente, se obtiene una velocidad de propagación distinta:

[177] Ver el capítulo 13 del libro de P. Bergmann "*Introduction to the theory of relativity*", Prentice-Hall, New York, 1942.

$$\left| r \cdot \frac{d\theta}{dt} \right| = \sqrt{\left(1 - \frac{r_S}{r}\right)} \cdot c$$

Esta anisotropía en la propagación de la luz en un campo gravitatorio es solo aparente, debido a que la componente radial se calcula a partir de una coordenada que por sí misma no posee significado métrico.

Es posible efectuar un cambio de coordenadas tal que la velocidad de la luz sea la misma en todas direcciones (coordenadas isótropas).

Las coordenadas isótropas se obtienen directamente de la expresión del elemento espacial del intervalo: $dl^2 = \frac{dr^2}{\left(1 - \frac{r_S}{r}\right)} + r^2 \cdot d\theta^2 + r^2 \cdot sen^2\theta \cdot d\phi^2 = dl_r^2 + dl_\theta^2 + dl_\varphi^2$

Expresada a partir del intervalo, la velocidad radial es $\frac{dl_r}{dt} = \frac{\frac{dr}{dt}}{\sqrt{1 - \frac{r_S}{r}}} = \sqrt{\left(1 - \frac{r_S}{r}\right)} \cdot c$.

Entonces, la velocidad radial de un corpúsculo luminoso $\frac{dl_r}{dt}$ resulta igual a las velocidades tangenciales $\frac{dl_\theta}{dt}$ y $\frac{dl_\varphi}{dt}$. Pero el efecto que siempre permanece **es el enlentecimiento de la luz en las proximidades de las masas gravitantes**.

En un campo gravitatorio el postulado de la relatividad restringida, según el cual la velocidad de la luz en el vacío es constantemente igual a la constante universal c, **ya no se cumple** en general. No obstante, también en la teoría generalizada la velocidad de la luz es la misma, igual a c en cualquier sitio y en cualquier dirección respecto de un marco inercial local. Precisamente eso es lo que significan tanto la anulación del intervalo cuando se la interpreta desde un marco inercial de referencia como el límite común c de las velocidades radial y tangencial cuando la coordenada radial r de la métrica de Schwarzschild tiende a infinito.

11.3 Corrimiento al rojo en un campo gravitatorio

Consideremos una masa central en cuyo campo gravitatorio se encuentra un observador a una distancia r de la masa.

La métrica de Schwarzschild define una variedad de Riemann a través de las coordenadas $x_0 = c \cdot t \quad x_1 = r \quad x_2 = \theta \quad x_3 = \phi$ y el tensor métrico de componentes

$$g_{00} = 1 - \frac{r_S}{r} \qquad g_{11} = -\frac{1}{1 - \frac{r_S}{r}} \qquad g_{22} = -r \qquad g_{33} = -r \cdot sen\theta \qquad g_{i,k} = 0 \quad i \neq k$$

Construimos el hiper-plano tangente (un espacio de Minkowski) por el punto de la variedad donde se encuentra el observador. Ese espacio tangente de Minkowski se encuentra libre de efectos gravitatorios: es plano. Sus coordenadas son $X_0 = c \cdot T \quad X_1 \quad X_2 \quad X_3$. El origen del espacio tangente coincide con el punto de tangencia en la variedad curvada.

El elemento de línea en el punto es el mismo para la variedad plana tangente y para la variedad curvada:

$$ds^2 = \left(1 - \frac{r_S}{r}\right) \cdot c^2 \cdot dt^2 - \frac{dr^2}{\left(1 - \frac{r_S}{r}\right)} - r^2 \cdot d\theta^2 - r^2 \cdot sen^2\theta \cdot d\phi^2 = c^2 \cdot dT^2 - dX_1^2 - dX_2^2 - dX_3^2$$

Para una partícula en reposo en el punto del campo gravitatorio considerado:

$ds^2 = \left(1 - \frac{r_S}{r}\right) \cdot c^2 \cdot dt^2 = c^2 \cdot dT^2$ Entonces las medidas del tiempo son diferentes en el marco inercial libre de gravitación (espacio de Minkowski tangente) dT y en el marco de referencia del campo gravitatorio dt : $dT = \sqrt{1 - \frac{r_S}{r}} \cdot dt$

La misma relación se aplica a los períodos de una oscilación medidos en un mismo punto, en un marco libre de gravedad y en el marco del campo gravitatorio.

La relación correspondiente a sus magnitudes recíprocas, las frecuencias f_0 y f :

$f = \sqrt{1 - \frac{r_S}{r}} \cdot f_0$ Puesto que f es siempre menor que f_0, un campo gravitatorio produce siempre un corrimiento al rojo. **Este es un efecto exclusivamente local**.

Pero ese efecto local para una oscilación de la misma frecuencia f_0 (respecto de un marco inercial), produce un corrimiento al rojo diferente en puntos situados a distancias diferentes de la masa que crea el campo. Si una fuente en reposo situada a una distancia r_1 de la masa gravitante emite un pulso luminoso de frecuencia f_1 medida en el marco de referencia del campo, un observador situado a una distancia r_2 en el mismo campo gravitatorio cuando mide la frecuencia de ese pulso encuentra que es f_2.

Pero la fórmula $f = \sqrt{1 - \frac{r_S}{r}} \cdot f_0$ se aplica en ambos casos por lo cual: $\frac{f_2}{f_1} = \frac{\sqrt{1 - \frac{r_S}{r_2}}}{\sqrt{1 - \frac{r_S}{r_1}}}$

Como las longitudes de onda λ son inversamente proporcionales a las frecuencias, $\lambda = \frac{1}{\sqrt{1 - \frac{r_S}{r}}} \cdot \lambda_0$ En consecuencia, entre la longitud de onda en el sitio donde se halla la fuente de luz y la longitud de onda en el sitio donde se encuentra un receptor se verifica:

$$\frac{\lambda_2}{\lambda_1} = \frac{\sqrt{1 - \frac{r_S}{r_1}}}{\sqrt{1 - \frac{r_S}{r_2}}}$$

Recordemos el radio de Schwarzschild: $r_S = 2 \cdot \frac{G \cdot M}{c^2}$

Si se define un **corrimiento espectral** $C = \frac{\lambda_2 - \lambda_1}{\lambda_1} = \frac{f_1 - f_2}{f_2}$ se obtiene: $1 + C = \frac{\sqrt{1 - \frac{r_S}{r_1}}}{\sqrt{1 - \frac{r_S}{r_2}}}$

11.4 Movimiento de corpúsculos en un tiempo-espacio con la métrica de Schwarzschild: avance del perihelio de una órbita y deflexión de la dirección de propagación de la luz.

Consideremos ahora la ecuación de movimiento de un corpúsculo. Como vimos en 10.2.4, si se trata de una partícula, es decir, si posee masa de reposo no nula, se mueve siguiendo una geodésica temporaloide, mientras que si su masa de reposo es nula se mueve siguiendo una geodésica nula.
Cuando la masa de la partícula es lo bastante pequeña como para no perturbar la geometría del tiempo-espacio, es posible hallar una geodésica temporaloide que describe una órbita plana en torno a la única masa significativa, generadora del campo. Este caso puede utilizarse para describir el movimiento de un planeta en torno del Sol. La órbita de un planeta en torno al Sol que predice la mecánica newtoniana cuando se desprecia los efectos gravitatorios de los demás planetas y otros cuerpos materiales, es una elipse perfecta con el Sol en uno de los focos de la elipse. Esa elipse posee un tamaño y una forma especificados por el semieje mayor y la excentricidad (o en forma equivalente, por el semieje menor) y se encuentra fija en un plano denominado plano orbital. La orientación de la elipse en su plano puede determinarse por la dirección y sentido de la línea que une el sol con el planeta en el punto de la órbita en el que es mínima la distancia entre ambos, conocido como perihelio. De acuerdo con el modelo newtoniano de órbita elíptica, esta dirección no debería cambiar. De acuerdo con el modelo relativista la órbita no es una elipse perfecta: el perihelio avanza con una velocidad determinada por la teoría.
Como veremos, la geodésica nula puede deducirse a partir de la geodésica temporaloide efectuando un paso al límite, y permite predecir el ángulo en el que varía la dirección de un haz de luz cuando se aproxima desde el infinito a la masa generadora del campo, pero sin dirigirse directamente hacia ella, y después se aleja nuevamente al infinito.
Comencemos por deducir una ecuación para la geodésica temporaloide. Lo podemos hacer directamente a partir del intervalo, sin necesidad de calcular los coeficientes de conexión (las componentes del campo gravitatorio como las denominaba Einstein), empleando para ello algunos resultados elementales pertenecientes al cálculo de variaciones.

11.4.1 Elementos de cálculo de variaciones

Partimos de cuatro funciones regulares $x^0(u), x^1(u), x^2(u), x^3(u)$ que definen un arco de curva en el tiempo-espacio. El parámetro recorre el intervalo $[u_A, u_B]$, siendo A y B los puntos extremos del arco en ese continuo tetra-dimensional.
Por definición, para una curva temporaloide en el tiempo-espacio, el elemento de intervalo es definido positivo y se verifica:

$$\left(\frac{ds}{du}\right)^2 = \sum_{i,k=0}^{3} g_{i,k}\left(x^0(u), x^1(u), x^2(u), x^3(u)\right)\cdot \frac{dx^i(u)}{du}\cdot\frac{dx^k(u)}{du} > 0$$

Resumiendo con $\vec{x}(u)$ el conjunto formado $x^0(u), x^1(u), x^2(u), x^3(u)$, y con $\dot{\vec{x}}(u)$ el conjunto formado por sus derivadas respecto del parámetro, introducimos la función

$$F\left(\vec{x}(u), \dot{\vec{x}}(u)\right) = \sqrt{\sum_{i,k=0}^{3} g_{i,k}\left(x^0(u), x^1(u), x^2(u), x^3(u)\right)\cdot \frac{dx^i(u)}{du}\cdot\frac{dx^k(u)}{du}}$$

La longitud de arco de curva entre los puntos A y B viene dada por:

$$s_{A,B}[\vec{x}] = \int_{u_A}^{u_B} F\left(\vec{x}(u), \dot{\vec{x}}(u)\right).du$$

Una curva temporaloide que conecta dos puntos dados A y B, lo bastante próximos, es una geodésica si la longitud del arco correspondiente es mínima respecto de las longitudes que tienen los arcos de otras curvas vecinas que pasan por los mismos puntos extremos.

Entonces, hallar una geodésica temporaloide que une dos puntos A y B del continuo tetra-dimensional equivale a hallar una función $\vec{x}(u)$ que minimiza la integral que da la longitud de un arco de curva que conecta esos dos puntos.

Con el propósito de simplificar la exposición, trabajaremos con una sola función $x(u)$ y su derivada. Los resultados se extienden sin dificultad al caso de más de una función y más de una derivada.

Partamos de una función $x(u)$ y busquemos en qué condiciones hace que la integral tome un valor extremo en comparación con las funciones que toman los mismos valores que $x(u)$ en los extremos del intervalo de integración. A esas funciones las denominaremos funciones admisibles.

Si $\eta(u)$ es una función lo bastante regular, tal que se anula en el valor inicial u_A y final u_B del intervalo de valores del parámetro sobre el cual se integra, y si ε es un número que puede hacerse arbitrariamente pequeño, llegado el caso, entonces $x(u) + \varepsilon.\eta(u)$ es una función admisible, y todas ellas se pueden representar de esta forma.

Definamos la función $J(\varepsilon) = s_{A,B}[x + \varepsilon.\eta]$ y desarrollémosla en torno a $\varepsilon = 0$:

$$J(\varepsilon) = J(0) + \frac{dJ}{d\varepsilon}(0).\varepsilon + \frac{1}{2}\frac{d^2}{d\varepsilon^2}J(0)\cdot\varepsilon^2 + o\left(\varepsilon^2\right)$$

El término $o\left(\varepsilon^2\right)$ es un infinitésimo respecto de ε^2.

Para que $J(\varepsilon)$ presente un mínimo cuando $\varepsilon = 0$, y para que ese mínimo corresponda a la función $x(u)$, **es necesario** que se anule $\frac{dJ(0)}{d\varepsilon}$ para todas las funciones η concebibles que se anulen en los extremos del intervalo sobre el cual se integra. Definiendo $\dot{x} = \frac{dx}{du}$ y $\dot{\eta} = \frac{d\eta}{du}$ un cálculo inmediato arroja que:

$$\frac{dJ(0)}{d\varepsilon} = \int_{u_A}^{u_B} [\frac{\partial}{\partial x}F\left(x(u), \dot{x}(u)\right).\eta(u) + \frac{\partial}{\partial \dot{x}}F\left(x(u), \dot{x}(u)\right)\cdot\dot{\eta}(u)]\cdot du$$

Integrando por partes el término que presenta el factor $\dot{\eta}$ y teniendo en cuenta que η se anula en los extremos del intervalo de integración, se obtiene:

$$\frac{dJ(0)}{d\varepsilon} = \int_{u_A}^{u_B} [\left(\frac{\partial}{\partial x}F\left(x(u), \dot{x}(u)\right) - \frac{d}{du}\frac{\partial}{\partial \dot{x}}F\left(x(u), \dot{x}(u)\right)\right)\cdot\eta(u)]\cdot du$$

Como esta integral se debe anular para toda función $\eta(u)$, se puede aplicar el **lema de du-Bois Reymond**: si $\int_a^b g(u).\eta(u).du$, siendo $g(u)$ cualquier función continua en $[a,b]$, se anula para toda función $\eta(u)$, entonces $g(u) = 0$ para todo u perteneciente al intervalo de integración.

De este lema concluimos que la curva de longitud extrema $x(u)$ verifica la ecuación diferencial conocida como **ecuación de Euler**: $\frac{\partial F}{\partial x} - \frac{d}{du}\frac{\partial F}{\partial \dot{x}} = 0$

La solución de esta ecuación diferencial corresponde a un valor extremo de la integral. Para analizar si efectivamente se trata de un mínimo en general hay que estudiar los términos siguientes del desarrollo, a partir de $\frac{1}{2}\frac{d^2}{d\varepsilon^2}J(0)\cdot\varepsilon^2$, para todas las funciones admisibles. En el caso de la longitud de arco entre dos puntos lo bastante próximos la solución de la ecuación de Euler corresponde a un mínimo de la integral y por tanto determina la geodésica que buscamos.

Finalmente, un último resultado que utilizaremos en el estudio del movimiento geodésico: para toda curva $x(u)$ que verifica la ecuación de Euler con una función $F(x,\dot{x})$, que no depende en forma explícita del parámetro u, se deduce que $\frac{d}{du}\left(\dot{x}\cdot\frac{\partial F}{\partial \dot{x}} - F\right) = 0$ o sea que $\dot{x}\cdot\frac{\partial F}{\partial \dot{x}} - F = K$ es un **invariante** (es constante) sobre la curva considerada.

Consideremos ahora curvas $\vec{x}(u)$ en un espacio de n dimensiones y una función regular $F(\vec{x}(u),\dot{\vec{x}}(u))$ que por lo demás es una función cualquiera, no necesariamente restringida a la función que permite calcular una longitud de arco sobre la curva.
A cada función $\vec{x}(u)$ se le puede hacer corresponder un número real $I[\vec{x}]$ mediante la integral $I[\vec{x}] = \int_{u_A}^{u_B} F(\vec{x}(u),\dot{\vec{x}}(u))du$. La correspondencia $\vec{x}(u) \to I[\vec{x}]$ se denomina funcional. La longitud de un arco de curva es un ejemplo de funcional.
Los resultados anteriores se pueden extender al problema de hallar un valor extremo de la integral $I[\vec{x}] = \int_{u_A}^{u_B} F(\vec{x}(u),\dot{\vec{x}}(u))du$, es decir un extremo de la funciona $\vec{x}(u) \to I[\vec{x}]$.
Cuando $F(\vec{x},\dot{\vec{x}})$ depende de varias funciones coordenadas y de sus derivadas, obtenemos **una ecuación de Euler por cada coordenada** x^j, $\frac{\partial F}{\partial x^j} - \frac{d}{du}\frac{\partial F}{\partial \dot{x}^j} = 0$ y un **invariante escalar** $\sum_j \dot{x}^j \cdot \frac{\partial F}{\partial \dot{x}^j} - F = K$ sobre la curva. La suma se extiende sobre la totalidad de las coordenadas de la curva considerada y K varía de una curva a otra adyacente.

En general la ecuación de Euler da una condición necesaria para hallar la curva que minimiza la integral, pero no una condición de por sí suficiente. No obstante, cuando se trata de la longitud de un arco de curva regular n-dimensional que une dos puntos lo bastante próximos, la condición es también suficiente.

11.4.2 Geodésica temporaloide

Si tomamos el ángulo ϕ como parámetro, y expresamos la coordenada temporal y las otras dos coordenadas espaciales en función de este ángulo, entonces a partir de la métrica de Schwarzschild obtenemos para la longitud de arco de la geodésica que une los puntos A y B: $s_{A,B} = \int_{\phi_A}^{\phi_B} \frac{ds}{d\phi} \cdot d\phi = \int_{\phi_A}^{\phi_B} F\left(t(\phi), r(\phi), \theta(\phi), \dot{t}(\phi), \dot{r}(\phi), \dot{\theta}(\phi)\right) d\phi$

Por definición:

$$\left(\frac{ds}{d\phi}\right)^2 = F^2\left(t, r, \theta, \dot{t}, \dot{r}, \dot{\theta}\right) = \left(1 - \frac{r_S}{r}\right) \cdot c^2 \cdot \left(\frac{dt}{d\phi}\right)^2 - \frac{1}{\left(1 - \frac{r_S}{r}\right)} \cdot \left(\frac{dr}{d\phi}\right)^2 - r^2 \cdot \left(\frac{d\theta}{d\phi}\right)^2 - r^2 \cdot sen^2\theta$$

De acuerdo con los resultados del cálculo de variaciones presentados en 11.4.1, las funciones $t(\phi), r(\phi), \theta(\phi)$ que corresponden a una geodésica deben verificar las ecuaciones de Euler. Para $\theta(\phi)$: $\frac{\partial F}{\partial \theta} - \frac{d}{d\phi}\frac{\partial F}{\partial \dot{\theta}} = 0$ Calculando a partir de la fórmula para $F^2\left(t, r, \theta, \dot{t}, \dot{r}, \dot{\theta}\right)$, la ecuación de Euler para θ es: $\frac{d}{d\phi}\left(\frac{r^2 \cdot \dot{\theta}}{F}\right) = \frac{r^2 \cdot sen\theta \cdot \cos\theta}{F}$

Esta ecuación posee dos soluciones con $\theta(\phi)$ constante: $\theta(\phi) = 0$ y $\theta(\phi) = \frac{\pi}{2}$

La primera restringe la trayectoria a un movimiento a lo largo del eje z. No se trata de una órbita sino de un movimiento radial (a lo largo del eje z) desde o hacia la masa generadora del campo, que se encuentra en el origen de las coordenadas espaciales. Pero la otra solución constante $\theta(\phi) = \frac{\pi}{2}$ corresponde a un movimiento que se produce en el plano x-y y por tanto permite describir una órbita. Supongamos entonces que $\theta(\phi) = \frac{\pi}{2}$.

La ecuación de Euler para $t(\phi)$, $\frac{\partial F}{\partial t} - \frac{d}{d\phi}\frac{\partial F}{\partial \dot{t}} = 0$ es: $\frac{d}{d\phi}\left(\frac{c^2 \cdot \left(1 - \frac{r_S}{r}\right)}{F} \cdot \dot{t}\right) = 0$

Entonces la expresión entre paréntesis es una constante H y despejando $\dot{t} = \frac{dt}{d\phi}$ resulta:

$\dot{t} = \frac{H \cdot F}{c^2 \cdot \left(1 - \frac{r_S}{r}\right)}$ En vez de hallar la ecuación de Euler para r, es conveniente trabajar con el invariante escalar, que sabemos por lo expuesto en 11.4.1 que permanece constante sobre la trayectoria: $\dot{t} \cdot \frac{\partial F}{\partial \dot{t}} + \dot{r} \cdot \frac{\partial F}{\partial \dot{r}} + \dot{\theta} \cdot \frac{\partial F}{\partial \dot{\theta}} - F = K$

Esta última ecuación se puede re-escribir así, luego de multiplicar ambos miembros de la igualdad por la función F: $\frac{1}{2} \cdot \left(\dot{t} \cdot \frac{\partial F^2}{\partial \dot{t}} + \dot{r} \cdot \frac{\partial F^2}{\partial \dot{r}} + \dot{\theta} \cdot \frac{\partial F^2}{\partial \dot{\theta}}\right) - F^2 = K \cdot F$

De la definición de F^2 se desprende que:

$$\frac{1}{2}\cdot\left(\dot{t}\cdot\frac{\partial F^2}{\partial \dot{t}}+\dot{r}\cdot\frac{\partial F^2}{\partial \dot{r}}+\dot{\theta}\cdot\frac{\partial F^2}{\partial \dot{\theta}}\right)=F^2+r^2\cdot sen^2\theta$$

Como en este caso $\theta(\phi)=\frac{\pi}{2}$ es constante, resulta finalmente la siguiente relación:

$$\frac{1}{2}\cdot\left(\dot{t}\cdot\frac{\partial F^2}{\partial \dot{t}}+\dot{r}\cdot\frac{\partial F^2}{\partial \dot{r}}\right)-F^2=F^2+r^2-F^2=r^2=K\cdot F$$

Entonces: $F\left(t(\phi),r(\phi),\frac{\pi}{2},\dot{t}(\phi),\dot{r}(\phi),0\right)=\frac{r^2}{K}$ Puesto que F es positiva, la constante K también debe serlo.

Para hallar una ecuación diferencial en términos de r y de $\frac{dr}{d\phi}$ solamente, se sustituye la fórmula obtenida para $\frac{dt}{d\phi}$ en $F^2=\left(\frac{r^2}{K}\right)^2=\left(1-\frac{r_S}{r}\right)\cdot c^2\cdot\left(\frac{dt}{d\phi}\right)^2-\frac{1}{\left(1-\frac{r_S}{r}\right)}\cdot\left(\frac{dr}{d\phi}\right)^2-r^2$

Después de varios cálculos, de introducir la nueva coordenada $w(\phi)=\frac{1}{r(\phi)}$ en sustitución de la coordenada radial, luego de derivar respecto de ϕ y de eliminar el factor común $\frac{dw}{d\phi}$ resultante, se obtiene la siguiente ecuación diferencial que describe una órbita en el plano x-y: $\frac{d^2w}{d\phi^2}+w=\frac{r_S}{2\cdot K^2}+\frac{3}{2}\cdot r_S\cdot w^2$

Cuando la coordenada $r(\phi)$ crece lo suficiente, esta ecuación no lineal se puede aproximar por la ecuación lineal de un oscilador armónico de amplitud $w(\phi)$ con término forzante constante: $\frac{d^2w}{d\phi^2}+w=\frac{r_S}{2\cdot K^2}$

Integrando esta ecuación se obtiene $w(\phi)=\frac{1}{r(\phi)}=\frac{r_S}{2\cdot K^2}+C\cdot\cos(\phi-\phi_0)$ siendo C y ϕ_0 constantes de integración que dependen de los valores iniciales de $w(\phi)$ y de $\frac{d}{d\phi}w(\phi)$.

En algún punto de la órbita de la partícula se debe presentar un valor mínimo de la coordenada $r(\phi)$. Esta condición se puede satisfacer eligiendo $\phi_0=0$ de tal modo que $w(0)=\frac{r_S}{2\cdot K^2}\cdot(1+e)$ sea máximo y por tanto $\frac{d}{d\phi}w(0)=0$. Entonces la ecuación de la órbita se puede escribir así: $w(\phi)=\frac{1}{r(\phi)}=\frac{r_S}{2\cdot K^2}\cdot(1+e\cdot\cos(\phi))$

Definiendo $\frac{1}{p}=\frac{r_S}{2\cdot K^2}$ y despejando $r(\phi)$ resulta la bien conocida ecuación de una cónica en coordenadas polares con centro en uno de los focos: $r(\phi)=\frac{p}{1+e\cdot\cos(\phi)}$

Este es el mismo resultado que se obtiene en mecánica newtoniana para el movimiento de una partícula en el campo gravitatorio producido por una masa localizada en $r = 0$.
El **perihelio** (la mínima distancia entre la masa generadora y un punto de la órbita) vale $r_{\min} = \dfrac{p}{1+e}$.
La **excentricidad** e de la órbita debe ser estrictamente positiva. Cuando e se encuentra comprendida entre 0 y 1, la órbita es una elipse y el movimiento de la partícula en la aproximación newtoniana es periódico. En este caso el **afelio** (la máxima distancia entre la masa generadora y un punto de la órbita) se alcanza para valores angulares desplazados en π radianes respecto de los ángulos en los que se alcanza el perihelio, y vale $r_{máx} = \dfrac{p}{1-e}$.
Si la excentricidad vale 1 la órbita es una parábola, y si es mayor que 1 la órbita coincide con la rama de hipérbola adyacente al foco.

Cuando el término no lineal debe ser tenido en cuenta, se puede construir una solución aproximada aplicando una variante del método de perturbaciones de Poincaré-Linstedt. Para ello se asume una solución tentativa de la forma[178]:

$$w(\phi) = \frac{1}{r(\phi)} = A + B \cdot \cos(\zeta \cdot \phi) + C \cdot \cos(2 \cdot \zeta \cdot \phi) + \ldots$$

Al orden cero de aproximación, se tiene la misma órbita que en la mecánica newtoniana, pero al primer orden de aproximación resulta:

$$w(\phi) = \frac{1}{r(\phi)} = \frac{r_S}{2 \cdot K^2} \cdot (1 + e \cdot \cos(\zeta \cdot \phi)) \qquad \zeta \approx \frac{1}{1 + \dfrac{3 \cdot r_S^2}{4 \cdot K^2}}$$

Cuando la órbita está acotada en su plano, la aproximación de orden cero es una elipse de semieje mayor a. En este caso: $2 \cdot a = r_{mín} + r_{máx} = \dfrac{p}{1+e} + \dfrac{p}{1-e} = \dfrac{2 \cdot p}{1-e^2}$

Puesto que de la definición de p se puede concluir que $K^2 = \dfrac{r_S}{2 \cdot p}$, y $p = a \cdot (1 - e^2)$ resulta $K^2 = \dfrac{r_S}{2 \cdot a \cdot (1-e^2)}$ De esta última fórmula se desprende $\zeta \approx \dfrac{1}{1 + \dfrac{3 \cdot r_S}{2 \cdot a \cdot (1-e^2)}}$

Al primer orden de aproximación las magnitudes del perihelio y el afelio de la órbita no se alteran, pero la dirección del perihelio rota lentamente debido al factor ζ que aparece en el argumento del coseno.
Para que $\zeta \cdot \phi$ sufra un incremento de 2π y por tanto la partícula se encuentre nuevamente en el perihelio si en él se encontraba antes del incremento, ϕ debe haber sufrido un incremento de $2\pi + \delta\phi$ tal que $2\pi = \zeta \cdot (2\pi + \delta\phi)$.
Entonces **el perihelio avanza un ángulo** $\delta\phi = 2\pi \cdot \left(\dfrac{1}{\zeta} - 1\right) \approx \dfrac{3\pi \cdot r_S}{a \cdot (1-e^2)}$ **en cada revolución.** La órbita, casi elíptica, no se cierra sobre sí misma: rota en torno al foco

[178] Sobre este método, ver por ejemplo el libro de P. Kahn "*Mathematical Methods for Scientists and Engineers: Linear and Nonlinear Systems*", Dover, N.Y., 2004. Su aplicación en la teoría relativista de la gravitación se puede encontrar en A. Sommerfeld "*Electrodynamics*", Academic Press, New York, 1952.

donde se encuentra la masa generadora del campo, formando en el plano x-y una figura conocida como **roseta**.

Como $r_S = 2 \cdot \frac{G \cdot M}{c^2}$ es constante, la magnitud angular $\delta\phi$ de este movimiento de precesión puede aumentar solamente si se modifican las características geométricas de la órbita elíptica: si el semieje mayor de la órbita disminuye o si la excentricidad aumenta.

Estas modificaciones acercan la partícula a la región donde el tiempo-espacio se encuentra más curvado, es decir, en las proximidades de la masa generadora del campo. Por este motivo, en el sistema solar el mejor candidato para contrastar la predicción de la teoría generalizada sobre la precesión del perihelio con los datos de observación es el planeta Mercurio.

11.4.3 Geodésica nula

Para una geodésica nula, como es el caso de la trayectoria de un rayo de luz, $F^2(t, r, \theta, \dot{t}, \dot{r}, \dot{\theta}) = 0$ Este caso se puede reducir al caso estudiado en 11.4.2 si K se hace tender a $+\infty$. Se obtiene entonces la ecuación no lineal: $\frac{d^2 w}{d\phi^2} + w = \frac{3}{2} \cdot r_S \cdot w^2$

A partir de la solución obtenida mediante el método de perturbaciones mencionado en 11.4.2, haciendo tender K a infinito, ζ tiende a 1 y es posible demostrar que la siguiente es una solución aproximada para la trayectoria de la luz en el plano x-y:

$$w(\phi) = \frac{1}{r(\phi)} = \frac{3 \cdot r_S}{4} \cdot B^2 + B \cdot \cos(\phi) - \frac{r_S}{4} \cdot B^2 \cdot \cos(2 \cdot \phi)$$

Supongamos que la mínima distancia de la trayectoria a la masa generadora del campo es R y se alcanza para $\phi = 0$.

Entonces $w(0) = \frac{1}{R} = \frac{3 \cdot r_S}{4} \cdot B^2 + B - \frac{r_S}{4} \cdot B^2 = B + \frac{r_S}{2} \cdot B^2$

Si R es mucho mayor que $r_S = 2 \cdot \frac{G \cdot M}{c^2}$, resulta que $B \cong \frac{1}{R}$. Teniendo en cuenta la relación entre coordenadas cartesianas y coordenadas polares $x = r \cdot \cos(\phi)$, $y = r \cdot sen(\phi)$ la trayectoria $w = \frac{1}{r} = \frac{3 \cdot r_S}{4} \cdot \frac{1}{R^2} + \frac{1}{R} \cdot \cos(\phi) - \frac{r_S}{4} \cdot \frac{1}{R^2} \cdot \cos(2 \cdot \phi)$ puede escribirse, luego de multiplicar por $r \cdot R$: $R = \frac{3 \cdot r_S}{4 \cdot R} \cdot \sqrt{x^2 + y^2} + x - \frac{r_S}{4 \cdot R} \cdot \frac{x^2 - y^2}{\sqrt{x^2 + y^2}}$

En el plano x-y la trayectoria dada por esta función implícita se parece a una hipérbola. Presenta dos asíntotas, una corresponde a la dirección por la cual los fotones se aproximan a la masa desde el infinito y la otra corresponde a la dirección a lo largo de la cual se alejan al infinito.

Esas asíntotas vienen dadas por las fórmulas $x = R \pm \frac{r_S}{R} \cdot y$

Puesto que la magnitud $\frac{r_S}{R}$ de las pendientes es muy pequeña, se pueden igualar a los ángulos que las asíntotas forman con el eje y.

El ángulo $\frac{2 \cdot r_S}{R}$ que forman las asíntotas da la deflexión de la trayectoria de la luz, debida al campo gravitatorio, respecto de su dirección original. Evidentemente, cuanto mayor sea $r_S = 2 \cdot \frac{G \cdot M}{c^2}$ y menor sea R, mayor será la deflexión. Por tanto, para maximizar el efecto, el cuerpo generador del campo debe ser lo más masivo que sea posible y la luz debe incidir rasante a su superficie. Por este motivo las primeras confrontaciones de la predicción relativista de la deflexión de la luz con los datos de observación se efectuaron empleando la luz de una estrella incidiendo rasante en la superficie del Sol, durante un eclipse de este astro.

11.5 Confirmaciones de la teoría de la gravitación de Einstein

Desde 1919 hasta 2003 se han efectuado un conjunto de observaciones y experimentos tendientes a contrastar algunas consecuencias cuantitativas de la teoría de la gravitación que se conviene en agrupar bajo la denominación de pruebas clásicas[179].
Estas se clasifican en cuatro categorías:

1. La precesión del perihelio de mercurio.
Como vimos en 11.4.2, según la mecánica newtoniana la dirección del perihelio de un planeta no debería cambiar. Pero ya en 1845 se había observado que con cada órbita sucesiva la dirección del perihelio de Mercurio en su plano orbital se modificaba ligeramente en un movimiento conocido desde entonces como precesión del perihelio.
Parte de este movimiento se pudo explicar por los efectos gravitatorios de los demás planetas, pero permaneció una rotación residual de 43” de arco cada 100 años, sin explicación satisfactoria en el marco de la mecánica clásica.
En 1915, a partir de una solución aproximada de las ecuaciones del campo, Einstein mostró que la nueva teoría de la gravitación permitía predecir el avance del perihelio. A menos de una incertidumbre del 1 %, la predicción basada en la teoría está de acuerdo con las mediciones. La teoría restringida conduce a una precesión del perihelio debido a la variación de la masa relativista con la velocidad de la partícula, pero predice la sexta parte del valor obtenido mediante la teoría generalizada.

2. La deflexión de la dirección de propagación de la luz de las estrellas por efecto del campo gravitatorio del Sol.
Este efecto ya había sido estudiado en el marco de la mecánica clásica, desde el punto de vista de la teoría corpuscular de la luz introducida por Newton. La deflexión calculada resulta ser la mitad de la deflexión relativista.
Cuando un rayo luminoso pasa rasante a la superficie del Sol, el acuerdo entre la teoría de la gravitación de Einstein y las mediciones para luz visible se produce a menos de un error del 10 %.
El acuerdo en el caso de microondas provenientes de quásares (fuentes de ondas electromagnéticas extremadamente lejanas y potentes que se consideran en el próximo capítulo), detectadas mediante el procedimiento conocido como VLBI (very long

[179] Una descripción bastante detallada de las pruebas ya efectuadas, o previstas dentro de los próximos años, puede hallarse en el libro de R. Lambourne “*Relativity, gravitation and cosmology*”, capítulo 7, Cambridge University Press, Cambridge, U.K., 2010. Se puede encontrar un resumen, a nivel más avanzado, al final del primer capítulo del libro de D. Ivanenko y G. Sardanashvili “Gravitación”, Editorial URSS, Moscú, 2005

baseline interferometry) que emplea dos o más antenas receptoras muy separadas (hasta miles de kilómetros), se produce a menos de un error del 0.04 %.

3. Corrimiento gravitacional hacia el rojo.
En este caso hubo que esperar hasta el experimento terrestre de Pound y Rebka (1960) para trabajar con buena precisión. Este experimento consiste en mediciones de laboratorio efectuadas sobre un fotón gamma que es emitido por un núcleo de una muestra emisora ubicada en lo alto de una torre, cae verticalmente 22.5 m en el campo gravitatorio terrestre, y es absorbido en la base por un núcleo de una muestra receptora. La implementación del experimento se efectuó utilizando el efecto Mössbauer (para lograr una frecuencia de emisión bien definida) y vibraciones mecánicas (para inducir corrimientos por efecto Doppler capaces de cancelar el corrimiento por efecto gravitatorio). El acuerdo entre la predicción teórica y los resultados experimentales se produjo a menos de un error del 10 %. Experimentos más refinados realizados posteriormente por Pound y Snyder mejoraron el acuerdo entre teoría y experimentos a menos del 1 %.
Poco después de formulada la teoría generalizada de la relatividad se comenzaron a obtener confirmaciones examinando el espectro de unas estrellas extraordinariamente densas, las enanas blancas. No obstante, la precisión de las pruebas efectuadas sobre una enana blanca descubierta en 1915, Sirio B, así como sobre otras enanas blancas, no se pudo incrementar en forma significativa hasta después de efectuado el experimento, en un laboratorio en la Tierra, por Pound y Rebka.
El corrimiento gravitacional al rojo fue verificado en 1979 por el equipo de A. Vessot, utilizando un máser de hidrógeno enviado a bordo de un cohete Scout, midiendo la variación en la frecuencia a medida que se modificaba el potencial gravitatorio en una órbita a 10 000 km de altura Este experimento verificó la dilatación del tiempo debida al campo gravitatorio de la Tierra a menos de un error de 0.007 %, por lo cual se considera que constituye una de las comprobaciones más precisas de la teoría general de la relatividad.
El funcionamiento continuado del sistema de posicionamiento global (GPS) también confirma la dilatación gravitatoria del tiempo a menos de un error diario del 1 %.

4. Retardo en los pulsos de radar debido al campo gravitatorio del Sol.
Las pruebas se han efectuado de la forma propuesta por Shapiro en 1964 o en forma modificada. El último experimento, efectuado utilizando la sonda espacial Cassini en 2003, arrojó un acuerdo entre la predicción teórica basada en la teoría de Einstein y las mediciones a menos de un 0.002% de error.

También se han efectuado experimentos basados en satélites y destinados a comparar las predicciones de la teoría con las mediciones realizadas en relación con la precesión de un giróscopo en dos circunstancias: cuando domina el efecto geodésico y cuando domina el efecto Lense-Thirring. Ambos son efectos extremadamente pequeños cuando se miden sobre una órbita en torno a la Tierra.
El efecto geodésico consiste en la precesión del eje de rotación de un giróscopo que se encuentra en una órbita polar en torno a una masa gravitante que no necesariamente rota respecto de un marco inercial local. El movimiento del eje de rotación es un giro en torno a la dirección radial, apuntando hacia el exterior de la masa, que adoptaría dicho eje en ausencia de este efecto.

El efecto Lense-Thirring, también conocido como efecto de arrastre de los marcos inerciales por un cuerpo masivo en rotación, se refleja en un efecto de precesión distinto del geodésico. Aparece como otra precesión en el eje de rotación de un giróscopo en órbita polar en torno a una masa gravitante que se encuentra en rotación respecto de un marco inercial local.
Los resultados obtenidos a partir de las mediciones efectuadas sobre los satélites LAGEOS I (puesto en órbita en 1976) y II (puesto en órbita en 1992) y formados por esferas de 0.6 m de diámetro y 411 kg, cubiertas con relectores para luz láser, no ponen en cuestión las predicciones de la teoría en relación con el efecto Lense-Thirring pero tampoco las confirman satisfactoriamente.
En 2004 se puso en órbita en torno a la Tierra la sonda de gravedad B, que permite detectar y eventualmente medir el efecto geodésico y el efecto de arrastre de marcos inerciales.
El acuerdo entre teoría y experimento en el caso del efecto geodésico se produce a menos de un 1.5 % de error, pero cuando el cuerpo masivo en rotación es la Tierra el efecto Lense-Thirring parece hallarse, por ahora, por debajo del nivel de ruido (o sea de las fluctuaciones aleatorias) asociado con las mediciones.

Finalmente, ciertas observaciones astronómicas sobre lentes gravitacionales (que se consideran en la próxima parte 11.6) y sobre sistemas gravitantes que se supone incluyen agujeros negros (que se consideran en la parte 13.3) se toman como evidencia indirecta a favor de la teoría de la gravitación de Einstein.

11.6 Óptica gravitacional: lentes, espejismos y materia oscura

A distancias lo bastante grandes de la masa gravitante, la métrica de Schwarzschild se puede aproximar así[180]:

$$ds^2 = \left(1 - \frac{r_S}{r}\right) \cdot c^2 \cdot dt^2 - \left(1 + \frac{r_S}{r}\right)\left(d\vec{r} \bullet d\vec{r}\right) \qquad d\vec{r} \bullet d\vec{r} = dr^2 + r^2 \cdot d\theta^2 + r^2 \cdot sen^2\theta \cdot d\varphi^2$$

El término $\left(1 \pm \frac{r_S}{r}\right)$ se puede reemplazar por $\left(1 \mp \frac{2 \cdot \Phi(r)}{c^2}\right)$, donde $\Phi(r)$ representa un potencial escalar newtoniano.

Puesto que para la trayectoria de un fotón el intervalo se anula, de la métrica aproximada se deduce la siguiente expresión para la velocidad del fotón en un potencial gravitatorio débil:

$$v = c \cdot \sqrt{\frac{1 + \frac{2 \cdot \Phi(r)}{c^2}}{1 - \frac{2 \cdot \Phi(r)}{c^2}}} \approx c \cdot \left(1 + \frac{2 \cdot \Phi(r)}{c^2}\right)$$

A partir de esta fórmula se puede calcular el tiempo que le insume al fotón ir del punto A al punto B del espacio físico: $T_{A \to B} = \int_A^B \frac{dl}{v}$

La integración se hace respecto de un elemento de longitud dl de trayectoria espacial.

[180] L. Landau y E. Lifchitz, "*Teoría del campo*", Reverté, Barcelona, 1982, § 97.

Una aplicación del Principio de Fermat del tiempo extremo permite hallar la trayectoria seguida por el fotón en su viaje a través del campo gravitatorio.[181]
Las variaciones en el potencial gravitatorio equivalen a modificaciones en el índice de refracción de un medio material: una verdadera óptica gravitacional se hace posible gracias a la teoría de la relatividad.

Cuando una o más masas galácticas, o las masas ocultas asociadas a la denominada materia negra del universo, se interponen entre una fuente de radiación electromagnética del observador, se comportan como verdaderas lentes gravitatorias.
La Figura 11.1 muestra un esquema de este efecto actuando sobre la radiación emitida por un quásar que llega a nuestro planeta en forma directa y también luego de haber sufrido una deflexión en el campo gravitatorio de una galaxia u otro cuerpo muy masivo.

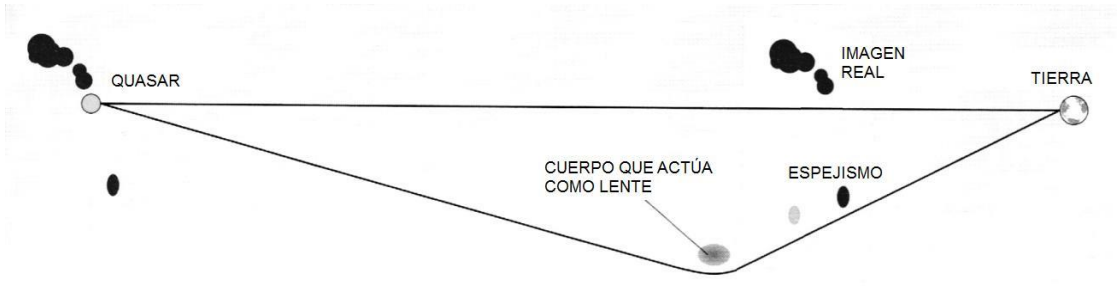


Figura 11.1. Efecto de lente gravitatoria

La primera observación de un espejismo gravitacional, es decir, de la creación de imágenes múltiples de una misma fuente por una masa interpuesta entre la fuente y el observador (en este caso situado en la Tierra) parece haber sido el "quásar doble" producido por la lente gravitacional formada por una galaxia y reportado en 1979[182].

Utilizando métodos computacionales de procesamiento de imágenes, resulta posible suprimir las imágenes del quásar de modo de hacer aparecer una imagen de la galaxia que actúa como lente gravitacional.
Lo mismo se puede hacer en el caso de la Cruz de Einstein, que consiste en la imagen cuádruple de un pulsar (Q2237+0305) producida por una galaxia espiral, mostrado en la Figura 11.2.

[181] Sobre el principio de Fermat y otros principios de extremo, incluidas aplicaciones a lentes gravitacionales, puede verse el libro de J.L. Basdevant "*Variational principles in Physics*", Springer Verlag, New York, 2007.

[182] Los quásares y los pulsares son fuentes de radiación electromagnética cósmica cuyas propiedades son bastante enigmáticas. Los quásares emiten energía a un ritmo miles de millones de veces mayor que el del sol, mientras que los pulsares emiten una radiación de muy alta frecuencia en forma de pulsos, con una periodicidad de un segundo o menos. Sobre estos y otros enigmáticos objetos cósmicos puede verse el libro de J. Rich "*Fundamentals of Cosmology*", Springer Verlag, Heildelberg, 2001.

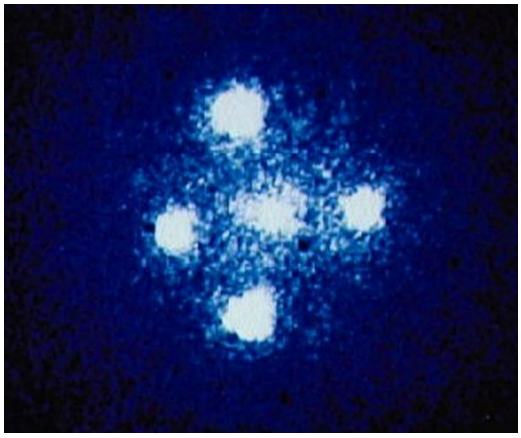


Figura 11.2. Cruz de Einstein

Finalmente, la óptica gravitatoria se conecta con el tema de la materia oscura: una forma de materia que se propuso para explicar ciertos efectos dinámicos, como los relacionados con la velocidad de rotación de las estrellas en las galaxias, con los movimientos de las galaxias en los cúmulos galácticos y con los movimientos de los cúmulos en agregados de orden superior.
La mecánica clásica se puede aplicar al estudio de la dinámica de los sistemas de cuerpos sometidos a interacciones gravitatorias, como las galaxias y cúmulos de galaxias.
Los cuerpos componentes de estos sistemas, fundamentalmente estrellas y polvo cósmico, interactúan fuertemente entre sí mediante interacciones gravitatorias locales que se pueden describir mediante la teoría de Newton.
Además interactúan con el resto de la materia del cosmos en forma más débil a través de la curvatura promedio del tiempo-espacio-materia en la región en la que esos sistemas se encuentran.
Debido a que predomina la interacción gravitatoria dentro del sistema en la determinación de su dinámica, se los suele denominar sistemas auto-gravitantes.

El teorema del virial[183] se puede aplicar a un sistema auto-gravitante que no se expande ni se contrae, es decir, tal que los cuerpos que lo componen no presentan una componente permanente de velocidad en dirección radial respecto del centro de masas, alejándose o acercándose del mismo.
En ese caso, este teorema afirma que: $2 \cdot E_c + \Phi = 0$
En esta ecuación E_c es la suma de las energías cinéticas de los N cuerpos que componen el sistema auto-gravitante y Φ es su energía potencial de interacción gravitatoria. Si $M = \sum_{k=1}^{N} m_k$ es la suma de las masas m_k de esos cuerpos, y $\langle v^2 \rangle = \frac{1}{M} \cdot \sum_{k=1}^{N} m_k \cdot v_k^2$ es el promedio del cuadrado de las velocidades, entonces $2 \cdot E_c = M \cdot \langle v^2 \rangle$. Por otra parte, se puede estimar $\Phi \approx -\frac{G \cdot M^2}{R}$, siendo R una dimensión característica del sistema (el radio si se trata de una estrella).

[183] Ver por, ejemplo, L. Landau y E. Lifchitz "*Mecánica*", capítulo 2, Reverté, Barcelona, 1981.

Entonces, del teorema del virial se desprende la relación aproximada $\langle v^2 \rangle \approx \frac{G \cdot M}{R}$ que permite estimar una de las variables si se pueden determinar las otras dos.
Cuando se aplica esta relación a los sistemas de estrellas lo bastante extensos, para los cuales se pueden estimar R, M y $\langle v^2 \rangle$ en forma independiente, se encuentra que $\langle v^2 \rangle$ es significativamente mayor que $\frac{G \cdot M}{R}$.
Admitiendo que la estimación de R es correcta, entonces M debe haber sido subestimada. La masa faltante es lo que se conoce como **materia oscura**.

En la práctica se procede partiendo de datos de observación para estimar el valor promedio del cuadrado de la velocidad relativa entre pares de galaxias.
Este resultado junto con el teorema del virial de la mecánica clásica se utiliza entonces para estimar la profundidad del pozo de potencial gravitatorio promedio en la parte del Universo accesible a la observación. A su vez, a partir de este último dato se puede estimar la densidad promedio del universo.
Otros datos obtenidos a partir de observaciones permiten una estimación independiente de dicha densidad promedio, pero este último resultado da entre un 5 y un 10 % del valor obtenido partiendo de las velocidades relativas de las galaxias.

La diferencia se supone corresponde a la materia oscura, mucho más numerosa que la materia ordinaria que compone las estrellas, los planetas, los cometas, los asteroides y el polvo intergaláctico. Se la denomina materia oscura fría debido a que las estimaciones de las velocidades de las partículas que la componen dan resultados mucho menores que la velocidad de la luz.
Formando parte de la materia oscura galáctica, podrían existir bariones[184] (protones, neutrones e hiperones) ocultos en sentido de no emitir ni absorber radiación electromagnética en forma significativa.
Esta materia oscura bariónica podría alterar las trayectorias de los fotones hasta producir efectos observables, como los halos galácticos y otros fenómenos, cuyo estudio se ha intensificado a partir del momento en que se dispuso de instrumentos astronómicos avanzados, como el telescopio Hubble.

La evidencia observacional reciente sobre el corrimiento al rojo de la emisión electromagnética de la materia ordinaria de las galaxias (cuyos protones y neutrones representan apenas un 4% de la densidad de materia-energía cósmica) y su interpretación en el contexto del denominado modelo estándar lambda-materia oscura fría, sugiere que, además de la materia oscura fría (22 % del total de la densidad actual de materia-energía), existiría una forma de energía mayoritaria (74% del total de la densidad de materia-energía), conocida como energía oscura (Figura 11.3).

[184] La materia oscura bariónica formaría parte de objetos oscuros, como planetas grandes y agujeros negros. La cosmología astrofísica en su estado actual conduce a plantear la existencia de una materia oscura no bariónica. Sobre este y otros temas relacionados con la materia oscura y con la denominada energía oscura pueden consultarse los capítulos de la primera parte de la obra colectiva editada por G. Fraser "*The new physics in the twenty first century*", Cambridge University Press, U.K, 2009.

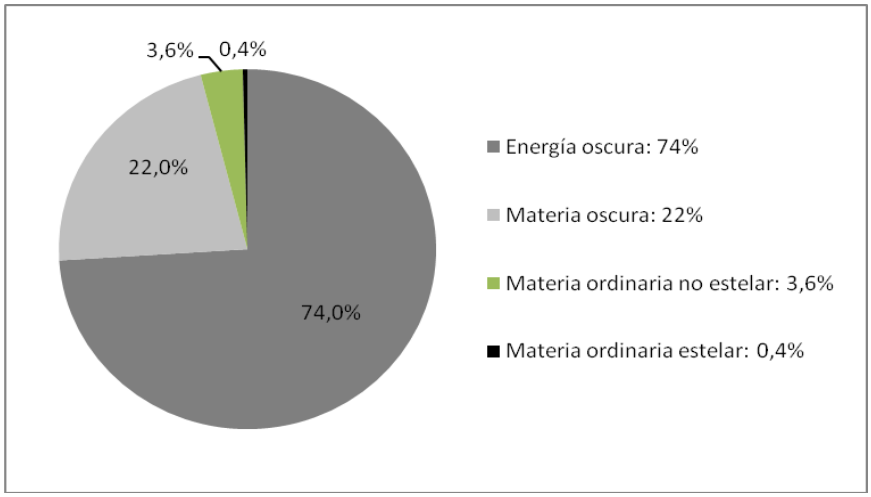


Figura 11.3 Distribución porcentual de las fuentes del campo gravitatorio tal como se las concibe en el momento actual. El aporte de la materia ordinaria es apenas un 4% del total.

Pero la materia oscura y la energía oscura nos conducen al próximo tema: el problema cosmológico.

12. El problema cosmológico

"Es una luz muy débil la que nos llega del cielo estrellado. ¿Qué sería, sin embargo, del pensamiento humano si nosotros no pudiéramos ver estas estrellas, como hubiera sucedido por ejemplo, si, al igual que su hermana Venus, la Tierra se envolviera siempre en un manto de nubes?"
Jean Perrin

"Cuanto menos uno sabe sobre el Universo, más fácil es explicarlo."
León Brunschvicg

"En 1915, la teoría general de la relatividad pretendía hacer participar a la naturaleza, de la libertad que el espíritu había reconquistado frente al espacio y reclamaba a la astronomía para que decidiera, por medio de muchas pruebas, si la geometría física, si la estructura del espacio-tiempo era euclidiana o no.
Tres fenómenos celestes estaban, a priori, dados como criterios; en pocos años la astronomía estuvo en condiciones de dar tres respuestas favorables a una geometría no euclidiana: la de Riemann… De este modo el espacio perdió todo carácter abstracto: se identifica con los campos de gravitación y tiene sus propiedades.
Un cuarto fenómeno astronómico… plantea la cuestión de la rigidez de los marcos de espacio-tiempo; se trata de la expansión del Universo, sobre el cual la observación y la teoría han enredado sus resultados entre 1925 y 1930. Lo esencial es el hecho siguiente: un espacio-tiempo estático se considera imposible*; el Universo tomado en su conjunto, no está en equilibrio, el espacio-tiempo aumenta su radio y la fuga de las galaxias refleja esta rápida evolución.*
Más, aquí, el pensamiento es traicionado por la indigencia del lenguaje vulgar. Originadas en un nivel inferior, las palabras son incapaces de expresar las relaciones donde, entretanto, se refugia la realidad del mundo; la verdad reside en las fórmulas donde se rompen las estrechas ligaduras de lo concreto, donde la ciencia escapa a la estrechez de las palabras…"
Paul Couderc

"La Astrofísica ha modificado todas nuestras ilusiones- en otras palabras, ha cambiado nuestra misma alma."
Anatole France

En el leguaje común, se considera que la cosmología estudia el Universo (literalmente, todo lo que existe) tal y como es ahora, mientras que la cosmogonía nos informa sobre cómo se formó el Universo. Pero si se tiene en cuenta que las ondas electromagnéticas que captamos hoy, con telescopios ópticos o radiotelescopios, han viajado con una velocidad finita, se advierte de inmediato que nos ofrecen imágenes de las galaxias tal como se encontraban en el pasado, inclusive en un pasado muy remoto.
Estamos captando información correspondiente a un intervalo de tiempo de miles de millones de años de vida del Universo: cosmología y cosmogonía ya no se pueden disociar.

La información procedente de nuestro grupo local de galaxias posiblemente corresponde, al menos a grandes rasgos, a su estado actual. La proveniente de otros cúmulos de galaxias corresponde al estado en el que se encontraban hace varios miles de millones de años, tiempo durante el cual estos sistemas pueden haber sufrido modificaciones profundas[185].
En una galaxia como la nuestra, se sabe ahora que hay estrellas que nacen, estrellas que explotan, estrellas que pulsan. Partículas de polvo cósmico y gases deambulan entre las estrellas y en el espacio cósmico intergaláctico. Se forman y se disgregan grupos de estrellas, a veces asociados a la rotación de las galaxias, de tal forma que cuando las galaxias están lo bastante próximas, pueden mezclar su materia estelar.
Las galaxias en traslación y rotación relativa se deforman y se dislocan, describiendo órbitas inmensas en los cúmulos a los que pertenecen.
Los cúmulos de galaxias forman cúmulos mayores y estos otros mayores aún, mientras que la totalidad parece estarse expandiendo con velocidad relativa tanto mayor cuanto más grande sea la distancia que separa una parte de otra.
Esta expansión del Universo ha constituido el problema principal de la Astronomía en el siglo XX. Su estudio teórico fue posible gracias a los modelos cosmológicos basados en la teoría generalizada de la relatividad.
Como vimos en el capítulo 11, la teoría generalizada de la relatividad tuvo un notable éxito en la explicación de fenómenos locales que se presentan en el Sistema Solar o (a escala galáctica) en sus adyacencias: el retardo de los relojes en los campos gravitatorios, la curvatura de los rayos de luz por el campo gravitatorio debido a masas estelares, y el movimiento de precesión de las órbitas planetarias. También, sirviendo de base a una verdadera óptica gravitatoria, como se vió en 11.6, permite explicar el efecto de lente gravitacional provocado por las galaxias.
Pero entre los fenómenos locales y la óptica gravitatoria por un lado, y por el otro el encare del problema cosmológico propiamente dicho, es decir, intentar responder a la interrogante acerca de cuál es la estructura espacio-temporal del Universo como un todo, su forma y su dinámica, hay una diferencia abismal.
Para poner en perspectiva esto último, conviene comenzar por revisar algunos aspectos históricos de la relación entre la cosmología, la cosmogonía y las ciencias físicas (incluyendo entre estas últimas a la Astronomía).

12.1 Aspectos históricos

Los primeros sistemas cosmológicos surgieron estrechamente vinculados con las religiones. En general ubicaron a la Tierra en el centro del cosmos, como una extensión plana apoyada en columnas o flotando en el agua, sobre la que se ubican los cielos, y por debajo de la cual se halla un inframundo de naturaleza infernal. Los aspectos cosmogónicos se planteaban en términos de una creación divina.
Los pensadores griegos supusieron que cada planeta, la Luna y el Sol se movían alrededor de la tierra en su correspondiente esfera celeste, una cáscara traslúcida. Todas estas esferas estaban contenidas en la esfera de las estrellas fijas, más allá de la cual no había nada, con excepción del espíritu divino. El origen de la idea de las esferas astrales se encuentra en la presuposición de que los cuerpos celestes solamente podían moverse

[185] La observación del cielo y la interpretación de los datos muestra que en las espiras de las galaxias se encuentran estrellas jóvenes, de edades del orden del millón de años, mientras que hacia el interior se hallan estrellas viejas, de edades del orden de los cinco mil millones de años.

con movimiento perfecto, circular uniforme, y que aparte de los cambios de posición debidos a este movimiento, ninguna otra modificación es posible en los cielos.
Cuando los fenómenos parecían contradecir las consecuencias de los principios, se buscaba una explicación para salvar los fenómenos. Si el Sol parecía presentar alguna variación en su aspecto visual, como una mancha en su superficie, se pensaba que ello era debido a una perturbación en la atmósfera que se interponía entre el astro y el observador. Si los movimientos aparentes de los planetas en la bóveda celeste parecían apartarse de un movimiento circular uniforme, ello se debía a que se combinaban varios de estos movimientos considerados perfectos. A medida que se fue perfeccionando la observación del movimiento de los planetas con el empleo de instrumentos más precisos, fue necesario ir aumentando el número de esferas hasta que a comienzos del siglo XVI se precisaron 80 esferas para dar cuenta de las observaciones.
Así pues, durante 2000 años el modelo geocéntrico con sus esferas celestiales dominó el pensamiento cosmológico[186], hasta la época de Copérnico, Galileo y Brahe y Kepler, cuando el centro del universo se trasladó de la Tierra al Sol, se registraron cambios en la supuestamente inmutable esfera de las estrellas y se comenzó a considerar como elipses a las órbitas de los planetas.

Resulta interesante analizar la razón por la cual la Astronomía no se desprendió, antes de fines del siglo XVI o comienzos del XVII, de la cosmovisión geocéntrica, pese a lo absurdo de las medidas destinadas a "salvar los fenómenos" en este marco tradicional y la tensión generada entre el viejo paradigma y los resultados de las observaciones.
Según Dingle[187], si la Astronomía hubiera sido un campo independiente de estudio, esta complejidad podría haber estimulado esfuerzos para reformarla mucho antes del momento en que efectivamente se reformó: "pero otras esferas de estudio estaban tan entrelazadas con la Astronomía que era imposible reformar la Astronomía sin reformar también la Física, la Química, la Fisiología, la Psicología y la Teología (por utilizar términos modernos para materias de estudio que por aquel entonces no se encontraban tan claramente diferenciadas como lo están ahora). No se hubiera podido hacer sin alterar la totalidad del esquema de creencias. El Universo era, en aquel tiempo, un Universo en un sentido mucho más literal de lo que lo ha sido desde entonces. Un diagrama tomado de un texto antiguo lo evidencia: el Cielo, con Dios, poseía una localización que era una parte del universo físico tanto como lo eran la tierra y las esferas cósmicas. Cada uno de los planetas tenía su influencia particular sobre el temperamento humano: así se establecieron nuestros adjetivos mercurial, marcial, jovial, saturnino. Una calamidad humana era un *desastre*- contra las estrellas. Una acción no natural era *exorbitada*, fuera de órbita. Los cuerpos terrestres se componían de cuatro elementos- tierra, agua, aire y fuego, cada uno de los cuales tendía a buscar "su propio lugar", y los cuerpos celestes estaban compuestos de un quinto elemento- una "*quintaesencia*" – que no tenía paralelo en la mudable Tierra… Era claramente muy peligroso introducir una modificación en el interior de este esquema cuyas partes se

[186] La idea de una Tierra esférica e incluso la idea de una Tierra que gira en torno al Sol aparecieron muy temprano en Grecia, pero fueron suplantadas por el sistema geocéntrico de Platón y Aristóteles.
La codificación del sistema geocéntrico, en el siglo II, D.C., por Claudio Ptolomeo facilitó su integración en una concepción del mundo que abarca toda la Edad Media y el Renacimiento.

[187] Herbert Dingle "Cosmology and science", en "*The Universe*", Simon & Schuster, New Yok, 1957, pp. 133-134.

encontraban tan íntimamente interrelacionadas, debido a los efectos imprevisibles sobre el todo."

Con la aparición de la mecánica newtoniana comienza un período de tres siglos durante los cuales la Astronomía se libera de la Cosmología. La mecánica celeste se centra en la dinámica del Sistema Solar. Los astrónomos dejan a los filósofos los interrogantes sobre el origen del universo, sobre si posee límites, y sobre hacia dónde se encamina. No intentan abarcar el Universo como un todo.
La Astronomía pre-relativista condujo a la creencia en un cosmos que ha existido siempre y que es infinito.
En él se encuentran astros iguales o parecidos al Sol y los planetas, moviéndose de acuerdo con las mismas leyes, y posiblemente distribuidos casi uniformemente en el espacio.
El espacio y el tiempo, independientes entre sí y absolutos, constituyen la forma de existencia de la materia y existen independientemente de ella.
Estas ideas, enseñadas durante tres siglos, caducaron en unos pocos años, entre 1920 y 1930, debido a la aparición de la teoría generalizada de la relatividad y a una notable secuencia de descubrimientos astronómicos.

A principios del siglo XIX, Wilhelm Olbers había planteado un problema que en el marco de un espacio infinito poblado de estrellas, con los conceptos de la mecánica newtoniana no tenía solución. Si las estrellas se distribuyen en forma uniforme a través del espacio, como parecía que lo hacían, entonces al mirar en cualquier dirección deberíamos encontrar una estrella más o menos cercana. Entonces el horizonte visual del cielo nocturno debería estar cubierto de luz, presentando a lo sumo algunos puntos negros, los planetas, que absorberían una parte de esa luz. En lugar de eso, el fondo del cielo nocturno es negro.
Otra dificultad, insalvable en un espacio infinito poblado de astros en todas partes, era una consecuencia de la ley de gravitación de Newton: el cálculo da que la fuerza de atracción sobre cualquier cuerpo material en estas condiciones debería ser infinita.
La salida a estas paradojas en el marco de la Física clásica fue descartar las investigaciones que trataban del Universo como un todo: de esta forma se podía ignorar las contradicciones.
La teoría de la relatividad hizo posible la construcción de modelos cosmológicos que superan esas contradicciones, empleando una nueva concepción de las relaciones entre tiempo, espacio y materia. Los problemas de la naturaleza, origen y destino del Universo volvieron a plantearse nuevamente, pero ahora en un marco científico y no en uno de filosofía especulativa.

12.2 El corrimiento del espectro de las galaxias y el horizonte de eventos

El análisis del espectro de la luz que proviene de una estrella permite identificar las líneas de algunos elementos químicos que deben hallarse en la atmósfera estelar y que también encontramos en la Tierra. Si la estrella se mueve radialmente respecto del observador, cabe esperar un corrimiento (efecto Doppler) de las líneas del elemento, respecto del espectro medido en reposo respecto del observador (en el laboratorio), hacia el rojo o hacia el azul según que la velocidad radial sea de alejamiento o de acercamiento al observador.

Recíprocamente, si observamos un corrimiento hacia el rojo de los espectros de la luz proveniente de una estrella de nuestra galarxia o de otra galaxia, podemos inferir que la estrella o la galaxia se alejan de nosotros. Esto es precisamente lo que encontraron Humason y Hubble para la casi totalidad de las galaxias[188].
Lo informaron en una publicación en 1929, en la cual calcularon la velocidad de recesión a partir de la fórmula clásica, no relativista, para el efecto Doppler.
Admitieron que la frecuencia ω' emitida por la galaxia se relaciona con la frecuencia ω recibida en la Tierra a través de la fórmula $\omega = \omega' \cdot \left(1 - \frac{V_r}{c}\right)$ donde V_r es la velocidad de recesión. Entonces, teniendo en cuenta que la longitud de onda λ es inversamente proporcional a la frecuencia: $V_r = c \cdot \left(\frac{\omega' - \omega}{\omega'}\right) = c \cdot \left(\frac{\lambda - \lambda'}{\lambda}\right)$

Reordenado: $\left(\frac{\lambda - \lambda'}{\lambda}\right) = \frac{\Delta\lambda}{\lambda} = \frac{V_r}{c}$ El corrimiento relativo al rojo $C = \frac{\Delta\lambda}{\lambda}$ es tanto mayor cuanto más alejada se encuentre la galaxia. Según la fórmula clásica, la velocidad radial es proporcional al corrimiento. De la relación entre el corrimiento C y las estimaciones de diatancia se desprende que las galaxias se alejan de nosotros con una velocidad radial V_r (generalmente dada en km/s) que parece ser proporcional a la distancia r (dada en años luz) que nos separa de esa galaxia: $V_r = H \cdot r$

La constante H se denomina constante de Hubble y $\frac{1}{H}$ posee dimensiones de tiempo[189].
Un observador situado en otra galaxia observaría el mismo fenómeno y enunciaría la ley de Hubble-Humason con la misma fórmula.

Generalizando esta ley se puede decir que las galaxias se alejan unas de otras a una velocidad proporcional a la distancia que las separa: al parecer el Universo se expande.
La ley de Hubble-Humason se puede re-escribir así: $\frac{1}{H} = \frac{r}{V_r}$

El miembro de la derecha en esta expresión se podría interpretar como el tiempo que le insume a una galaxia, desplazándose con velocidad uniforme V_r, para recorrer la distancia r. Como el miembro de la izquierda es el mismo para todas las galaxias, podría pensarse que todas las galaxias coincidieron en nuestra posición actual (o nosotros hubiéramos coincidido con ellas en un mismo punto): el Universo habría pasado por un estado donde toda la masa de las galaxias podría haber estado encerrada en la cabeza de un alfiler, por decirlo de alguna manera. Una especie de singularidad hiper-densa que luego se expandió explosivamente en una bola de fuego, con la radiación dominando la dinámica y la materia en forma de partículas relegada a un rol

[188] Solamente dentro del cúmulo de galaxias al que pertenece la Vía Láctea hay algunas galaxias que se aproximan a ella, como la nebulosa de Andrómeda. La totalidad de las galaxias pertenecientes a otros cúmulos se alejan.

[189] Las dimensiones de H son (km/s)/años luz. El año luz es una distancia (la recorrida por la luz moviéndose en el vacío durante un año).

secundario, parece algo fantástico, pero como se verá es compatible con la teoría generalizada de la relatividad.
Desde que fue estimado por vez primera, el valor numérico de la constante de Hubble ha sufrido varias modificaciones, siempre a la baja, a medida que se obtuvieron datos de objetos más lejanos y se pudo estimar las distancias de forma más precisa.
Actualmente se estima $\frac{1}{H}$ en un valor cercano a los 15.000 millones de años: esa sería una estimación grosera de la edad del Universo[190].

Si las galaxias se alejan unas de otras a una velocidad proporcional a su distancia, llega un momento en que su velocidad de recesión alcanza la velocidad de la luz.
Al final de 6.1 se hizo una distinción importante en relación con la adición relativista de velocidades. Por un lado se tiene el caso en el que un cuerpo se mueve respecto de un marco inercial K, el cual a su vez se mueve respecto de otro marco inercial K', y se busca la velocidad del cuerpo respecto de K': se aplica la ley de adición de velocidades relativista, de modo que la velocidad resultante no puede superar la velocidad de la luz.
Por otro lado se tiene el caso en el cual dos cuerpos A y B se mueven respecto del mismo marco inercial K, y se busca la velocidad relativa del cuerpo A respecto del cuerpo B: ahora se calcula según la fórmula de adición de velocidades clásica, porque todas las operaciones se realizan respecto del mismo marco de referencia.
Como consecuencia, si bien la velocidad c es un límite superior para todas las velocidades que pueden ser alcanzadas por los cuerpos materiales respecto de un marco de referencia inercial, la velocidad relativa de un cuerpo respecto de otro en el mismo marco de referencia puede resultar mayor que c.
Entonces dos cuerpos en movimiento respecto de un mismo sistema inercial en principio pueden presentar una velocidad de recesión mayor que c.
Pero de esto no se desprende que exista un cuerpo que pueda moverse respecto de K con esa velocidad mayor que c: no solamente en este caso no lo hay, sino que según la teoría de la relatividad no puede haberlo.
Entonces la velocidad relativa de dos cuerpos, determinada respecto del mismo marco de referencia inercial, no puede superar el doble de la velocidad de la luz.

El descubrimiento de los quásares (objetos casi-estelares al parecer mucho menores que una galaxia pero que radian la menos cien veces más energía que una galaxia grande y se encuentran en los confines del universo observable) en 1963, abrió un camino de investigación que condujo a una revisión de la ley de Hubble-Humason.
Resulta que Mc Vittie[191] y posteriormente otros investigadores de quásares hallaron corrimientos relativos $C = \frac{\Delta\lambda}{\lambda}$ bastante mayores que 3. Si el corrimiento es mayor que tres, la velocidad de recesión calculada mediante la fórmula clásica debe ser mayor que

[190] De acuerdo con la estimación más precisa disponible, basada en mediciones sobre la radiación de microondas cósmicas de fondo, esa edad sería de 13.73 ± 0.12 billones (10^9) de años. Estimaciones menos precisas arrojan edades comprendidas entre los 11 y los 20 billones de años.

[191] G. Mc Vittie "Distance and large redshifts" , The Quaterly Journal of the Royal Astronomical Society, 15 (3)

tres veces la velocidad de la luz, lo cual, según la teoría restringida no puede observarse respecto de un mismo marco inercial.
Pero cuando la velocidad de una fuente de luz se acerca a c es necesario utilizar la fórmula relativista deducida en 8.4.1 para conectar la frecuencia emitida en un marco inercial K' con la recibida por el observador fijo a otro marco inercial K:

$$\omega = \frac{\omega' \cdot \left(1 - \frac{V_r}{c}\right)}{\sqrt{1 - \frac{V_r^2}{c^2}}}$$

A partir de esta expresión resulta:

$$C = \frac{\Delta\lambda}{\lambda} = \frac{1 + \frac{V_r}{c}}{\sqrt{1 - \frac{V_r^2}{c^2}}} - 1$$

Entonces, admitiendo que es posible aplicar la teoría restringida de la relatividad, la ley de Hubble-Humason debería reformularse así:

$$\frac{1 + \frac{V_r}{c}}{\sqrt{1 - \frac{V_r^2}{c^2}}} - 1 = H \cdot r$$

En ese caso por grande que sea el corrimiento relativo, resulta siempre una velocidad radial inferior a c.

El problema es que en presencia de campo gravitatorio los marcos inerciales presentan un carácter local, mientras que los marcos inerciales de la relatividad restringida poseen un carácter global: cada uno abarca la totalidad del continuo tiempo-espacio.
Según la teoría generalizada, un tiempo espacio de Minkowski diferente debe emplearse en diferentes regiones: la variedad tetra-dimensional plana, tangente en cada punto de la variedad riemanniana formada por el tiempo-espacio-materia, ya no es la misma al pasar de un punto a otro.
Entonces un razonamiento basado en la teoría restringida de la relatividad queda invalidado porque en principio no es posible utilizar el mismo marco inercial para todo el Universo.

Si la velocidad de recesión se interpreta como una propiedad general del tiempo-espacio-materia es posible que la velocidad de recesión supere a la velocidad de la luz, dentro de ciertos límites, sin violar la teoría generalizada de la relatividad[192].
A escala cósmica tiempo y espacio pierden su sentido físico simple. Cuando Mc Vittie estimaba la velocidad de recesión del quásar a partir de una estimación de su distancia a la Tierra, la noción de distancia debería precisarse más. ¿Es la distancia entre el quásar y la Tierra cuando se emitió la luz o la distancia a la que se encuentra en el momento en el que el astrónomo efectúa la medición del corrimiento al rojo? ¿Continúa existiendo el quásar cuando el astrónomo lo percibe?
Como señala Paul Couderc[193]: "La velocidad calculada posee un significado complejo y analítico: no es ni la velocidad de la galaxia cuando parte el rayo luminoso, ni su velocidad en el instante de la percepción por el astrónomo. Además, en el espacio que

[192] Esto es una consecuencia del denominado modelo cosmológico estándar. Esta consecuencia se estudiará en la sección 12.4.3.

[193] P. Couderc "*L'Univers*" Colección ¿Que sais-je? Nº 687, Presses Universitaires de France, París, 1975.

se dilata llega un momento en que dos galaxias ya no pueden percibirse, pues el intervalo que las separa crece más deprisa que la velocidad de la luz."
Hay una separación a partir de la cual no llega información desde un objeto cósmico al otro: se ha alcanzado un horizonte de eventos, el **horizonte cosmológico**. Más allá el Universo nos resulta inaccesible, con independencia de las mejoras que podamos introducir en nuestros instrumentos y procedimientos de medición.

En todo caso, se advierte que las categorías del conocimiento humano, adecuadas para actuar en las escalas intermedias de tiempo y espacio de la experiencia cotidiana, e interpretables desde la perspectiva de la teoría de la evolución biológica como rasgos adaptativos, resultan inadecuadas para pensar el cosmos. Parecería que el único lenguaje utilizable es el lenguaje matemático. El único camino para interpretar y guiar las mediciones y comprender el Universo desde una perspectiva científica es construir modelos matemáticos del cosmos: tarea formidable pero al parecer ineludible.

12.3 Modelos de Universo

A medida que nos aproximamos a una masa la métrica del tiempo-espacio se modifica, poco o mucho según la magnitud de la masa y su concentración (densidad). Entonces, en las proximidades de una estrella la curvatura del tiempo-espacio varía. Esto se desprende de las ecuaciones del campo gravitatorio de Einstein: es un efecto local confirmado experimentalmente por observaciones astronómicas cuidadosas.
Una vez admitido que el continuo tiempo-espacio no es euclidiano, parecería que el problema de la forma del Universo considerado como un todo se podría plantear en forma parecida a como en la antigüedad se planteó el problema de la forma de la Tierra.
Dejando de lado argumentos de filosofía especulativa[194] y concentrando la atención en las observaciones y mediciones, se disponía de dos clases de evidencia sobre la forma de la Tierra.
Un ejemplo de evidencia del primer tipo es la aparición gradual de los barcos que se dirigen a tierra (comenzando por el extremo del mástil más alto). Este tipo de evidencia es local y débil, porque la curvatura de una parte comparativamente pequeña de la Tierra, por sí sola no asegura que la Tierra esté globalmente curvada y tenga una forma aproximadamente esférica.
Un ejemplo de evidencia del segundo tipo es forma circular de la sombra de la Tierra sobre la Luna durante los eclipses de Luna, en forma independiente del sitio en la Tierra desde el cual se observa el eclipse. Este tipo de evidencia es global y suministra un argumento empírico muy poderoso para afirmar la esfericidad de la Tierra, cuya primera confirmación preliminar se obtuvo luego de los numerosos viajes de circunnavegación, pero cuya confirmación definitiva se obtuvo gracias a los avances en astronáutica, a través de observaciones y mediciones en satélites artificiales y cápsulas espaciales.
Pero en cosmología toda la evidencia sobre la curvatura del Universo es local, corresponde a la evidencia débil, del primer tipo, sobre la esfericidad de la Tierra.
No se dispone de evidencia análoga a la evidencia de segundo tipo sobre la curvatura de la Tierra. La razón de esta carencia es que esta última evidencia se apoya en una tercera dimensión espacial, adicional a las dos dimensiones que caracterizan los puntos en la superficie de nuestro planeta, dimensión extra que en cosmología no existe.

[194] Como el argumento debido a Aristóteles: el movimiento natural de la materia pesada es dirigirse al centro del Universo, de modo que dicha materia se acumula allí dando origen, por razones de simetría, a una Tierra esférica.

Como nos encontramos confinados a una pequeña región del cosmos, para definir su topología global no podemos recurrir a otras evidencias que las locales, lo cual desde el inicio conduce a que el trabajo en torno al problema cosmológico sea de índole muy especulativa.
Admitido esto, podemos preguntarnos con Paul Couderc[195]:
"La materia que cubre el Universo, ¿es en conjunto incapaz de lograr, en gran escala, lo que cada grumo de realiza a su alrededor, en su escala? ¿Qué razones tenemos para postular *a priori* un Universo euclidiano e infinito, cuando comprobamos que alrededor de cada estrella, al menos, es no euclidiano?
Sin duda, el Universo tiene una estructura general menos simple que el marco euclidiano necesariamente infinito que la ciencia la atribuía sin escrúpulos, cuando no existían las geometrías no euclidianas.
Una vez surgida la duda, era necesario plantear el problema, considerar teóricamente todas las soluciones posibles y tratar luego de ver, mediante observaciones, a cuál de estas soluciones se adecúa el Universo (a condición de considerar más adelante las posibles complicaciones del problema, del mismo modo que se pasó de una Tierra esférica a un *geoide* muy complejo)."

La construcción de modelos cosmológicos se basa en tres suposiciones básicas que conviene explicitar.
En primer lugar, se supone que el Universo es un cosmos: posee estructura y dinámica a gran escala, sujetas a leyes.
En segundo lugar se supone que esa estructura y esa dinámica se pueden comprender a partir de las leyes físicas tal como las conocemos: se admite que esas leyes son universales pese a que solo las hemos podido confirmar en nuestros laboratorios y en el Sistema Solar[196].
Finalmente, se supone que la parte relativamente pequeña que hemos estudiado dentro de la parte del Universo que podemos observar, es una muestra representativa del conjunto[197].
Más concretamente, abordaremos el problema cosmológico desde la perspectiva física a partir de tres suposiciones básicas que, como se verá, se interrelacionan:
(1) El principio cosmológico en su versión débil.
(2) La aplicabilidad de las ecuaciones de Einstein para el campo gravitatorio al Universo en su totalidad.
(3) El postulado de Weyl.

---

[195] P. Couderc "*El Universo*", Eudeba, Buenos Aires, 1961, p. 59.

[196] Los progresos de la astronáutica, la instrumentación y las comunicaciones han posibilitado verificar a distancias astronómicas la estabilidad de los valores de varias constantes físicas fundamentales obtenidas en laboratorios de nuestro planeta. Ver, por ejemplo, E. Norman: "Are fundamental constants really constant?" American Journal of Physics, 54, 1986, pp. 317-321.

[197] Este último punto es particularmente delicado, debido a que cuando, por mecanismos de interacción se forman super-sistemas a partir de sistemas materiales, aparecen propiedades que restringen la dinámica que los sistemas poseen cuando no interactúan. Los astrónomos que estudian la estructura han encontrado evidencia sobre una estructura jerárquica hasta una escala de 10^9 años-luz, lo cual es ciertamente una distancia pequeña en relación con el denominado horizonte cosmológico que se mencionó en 12.2.

12.3.1 El principio cosmológico.

El Universo material conocido es un conjunto de estrellas y planetas, formando galaxias y cúmulos de galaxias, nubes cósmicas, gases y polvos interestelares y radiación. Si bien a una escala relativamente pequeña la distribución presenta irregularidades, se puede suponer como una primera aproximación que en gran escala esas irregularidades se desvanecen.

Con el término gran escala se hace referencia a dimensiones del orden de los 300×10^6 años-luz o más (1 año-luz=9.461×10^{15} m=0.307 parsecs) que se piensa caracterizan una gigantesca red de vacíos gigantes (estimados en un 90 % del volumen total) y super-cúmulos de galaxias (estimados en un 10 % del volumen total). [198]

Para escalas cuyo orden de magnitud se encuentra por debajo de los100×10^6 años-luz, el Universo muestra heterogeneidades y anisotropías, a medida que descendemos de los cúmulos de cúmulos de galaxias a los cúmulos de galaxias, a las galaxias por separado y aún más abajo.

Postulamos entonces que en gran escala el tiempo-espacio-materia se comporta como si estuviera formado por un fluido continuo, homogéneo e isótropo cuya densidad correspondería al valor que se obtendría distribuyendo uniformemente toda la energía de la radiación y las masas, incluyendo la proveniente de la materia oscura y la energía oscura.

A esta suposición le podemos llamar principio cosmológico débil[199]. Según el **principio cosmológico débil** el Universo presenta el mismo aspecto para todos los observadores situados en diferentes puntos del espacio.

El **principio cosmológico fuerte**, planteado por Bondi, Gold y Hoyle, agrega la suposición de que el Universo presenta el mismo aspecto para todos los observadores en cualquier sitio y en cualquier instante de tiempo pasado, presente o futuro.

La expansión del Universo condujo a los partidarios del principio cosmológico fuerte a postular una creación continua de materia. La estimación de la velocidad promedio de creación da del orden de un átomo por decímetro cúbico cada mil millones de años.

El Profesor Wheeler decía en forma algo irónica que el principio cosmológico fuerte y la hipótesis de la creación continua de materia dieron origen a la “era de la continua creación de teorías” en cosmología.

Mientras los principios cosmológicos se utilicen como hipótesis de trabajo, como generalizaciones útiles que pueden y deben ser abandonadas cuando entran en conflicto con los resultados de las observaciones, esos principios permiten el desarrollo de una cosmología de base relativista limitando lo que de otra forma sería un abanico inmanejable de posibilidades a tener en cuenta.

---

[198] Durante los últimos cincuenta años se ha venido acumulando evidencia experimental sobre la mencionada estructura de super-cúmulos de galaxias separados por vacíos gigantes. La evidencia más significativa al respecto se ha obtenido de un inventario detallado de quásares y del estudio del fondo de radiación cósmica fósil de micro-ondas. Sobre la evidencia observacional en la cual se sustenta el principio cosmológico así como el denominado **modelo cosmológico estándar**, puede consultarse el libro de S. Serjeant “*Observational cosmology*”, Cambridge University Press, U.K, 2010.

[199] Una presentación y una discusión muy interesantes de éste y otros principios cosmológicos, como los de Dirac y Milne, pueden encontrarse en la Parte II del libro de Jagit Singh “*Teorías de la cosmología moderna*”, Alianza, Madrid, 1974.

Lo que no resulta admisible es volver a ceder a la tentación de constreñir todos los fenómenos hasta acomodarlos en el marco de un principio cosmológico planteado a priori como cierto, en forma independiente de los fenómenos, como en general se hacía en las cosmologías de la Antigüedad y la Edad Media.

12.3.2 Las ecuaciones de la relatividad generalizada y la constante cosmológica

Cuando Einstein aplicó su teoría de la gravitación a la construcción de un modelo del cosmos, no se tenían evidencias de un Universo en expansión.
Las estrellas parecían moverse con velocidades pequeñas respecto de c, y en todas direcciones, por lo cual un modelo de Universo estático le pareció lo más adecuado. Supuso que la densidad promedio debía ser positiva y encontró que el problema cosmológico tenía una solución estática como la que buscaba, siempre que en cada punto del espacio existiera una presión negativa constante de magnitud igual a la mitad de la densidad promedio del Universo.
Como no había una justificación física conocida para esta acción local[200], Einstein modificó sus ecuaciones de la gravitación agregando un término, el término cosmológico $\Lambda \cdot g_{ik}$. Así resulta: $$R_{ik} - \frac{1}{2} \cdot g_{ik} \cdot R + \Lambda \cdot g_{ik} = -\kappa \cdot T_{ik}$$
La contante Λ es una constante, presumiblemente universal, pero introducida ad-hoc, conocida como **constante cosmológica**.
Las estimaciones de su orden de magnitud muestran que para el estudio de fenómenos a una escala local, como en el Sistema Solar y en la Vía Láctea, esa constante puede despreciarse.
Suponiendo que la constante cosmológica es positiva se puede deducir un modelo de Universo homogéneo, isótropo, finito y estático, compatible con la suposición de que existe en el espacio una densidad promedio de materia diferente de cero.
Einstein comenzó a arrepentirse de haberla introducido cuando actuó como árbitro del trabajo de Friedmann[201] en 1922 y constató que era posible resolver las ecuaciones del campo gravitatorio en presencia de una densidad promedio positiva de materia sin necesidad de introducir el término con la constante cosmológica, suponiendo que la solución no debe ser necesariamente estática. Una vez conocidos los resultados de Hubble y Humason sobre el corrimiento del espectro de las galaxias, Einstein abogó por su eliminación. Pero otros científicos eminentes, como Eddington, estaban convencidos de que la constante cosmológica no debía ser suprimida.

Observaciones astronómicas recientes[202], se interpretan como evidencia de que existe una gran cantidad de energía oscura que estaría acelerando la expansión del Universo.

---

[200] Al comienzo pensó que podia justificar un origen electromagnético para esa presión. Ver A. Einstein "*El significado de la relatividad*", 2ª edición, Espasa Calpe Argentina, Buenos Aires, 1952, Apéndice II.

[201] A. Friedmann "Über die Krümmung des Raumes",Z. Physik, 10, 1922, pp. 377-386.

[202] Ver, por ejemplo, el libro de S. Serjeant "*Observational cosmology*", Cambridge University Press, U.K, 2010, o la primera parte del libro editado por G. Fraser "*The new physics in the twenty first century*", Cambridge University Press, U.K, 2009.

Esto se puede tener en cuenta asignando un valor positivo adecuado a la constante cosmológica en el modelado matemático de la dinámica del Universo[203].
La energía oscura se puede tener en cuenta a través del tensor de energía impulsión, agregando una contribución de la forma $\overline{T}_{i,k} = \frac{\Lambda}{\kappa} \cdot g_{i,k}$, escribiendo las ecuaciones del campo así: $R_{i,k} - \frac{1}{2} \cdot R \cdot g_{i,k} = -\kappa \cdot \left(T_{i,k} + \overline{T}_{i,k}\right)$ Con el propósito de expresar el tensor de energía-impulsión en forma practicable para el desarrollo de modelos cosmológicos, se aplica el principio cosmológico y se supone que la materia, la radiación y la energía oscura pueden describir en gran escala mediante un fluido relativista, como el considerado en 8.7, pero ahora constituyendo el tiempo-espacio-materia cuando se prescinde de todas las heterogeneidades locales.
En ese caso cabría esperar que, si las componentes covariantes de la cuadrivelocidad del fluido oscuro se representan mediante U_i $(i = 0,1,2,3)$, la densidad de energía oscura por ρ_Λ y su aporte a la presión del fluido por p_Λ, entonces:

$$\overline{T}_{i,k} = \left(\rho_\Lambda \cdot c^2 + p_\Lambda\right) \cdot U_i \cdot U_k - p_\Lambda \cdot g_{i,k}$$

Si esta expresión debe ser idéntica a $\overline{T}_{i,k} = \frac{\Lambda}{\kappa} \cdot g_{i,k}$, la presión y la densidad del fluido oscuro deben verificar $p_\Lambda = -\frac{\Lambda}{\kappa}$ y $\rho_\Lambda = -\frac{p_\Lambda}{c^2} = \frac{\Lambda}{\kappa \cdot c^2}$. De este modo se anula idénticamente el término asociado a las componentes de la cuadrivelocidad, permaneciendo diferente de cero el término que expresa la presión del fluido oscuro.
Como una presión negativa corresponde a una tracción, el aporte de la energía oscura al fluido cósmico tiende a separar los objetos en el Universo a gran escala: si se quiere, es una especie de anti-gravedad.
Las ecuaciones del campo gravitatorio, teniendo en cuenta la energía oscura, adoptan la forma:

$$R_{i,k} - \frac{1}{2} \cdot R \cdot g_{i,k} = -\kappa \cdot \left(T_{i,k} + \rho_\Lambda \cdot c^2 \cdot g_{i,k}\right)$$

12.3.3 El flujo de Hubble y el postulado de Weyl.

El tiempo que aparece en las ecuaciones de la teoría generalizada de la relatividad es local y referido a un marco dado, pero en principio arbitrario.
El principio cosmológico débil plantea la extensión de la homogeneidad y la isotropía del espacio, tal como se las puede observar localmente efectuando un promedio a gran escala, a la totalidad del Universo.
Para ello es fundamental definir un marco de referencia local de modo que la velocidad media de la materia respecto de ese marco sea nula.
Ese marco puede considerarse como aproximadamente inercial, aunque solo sea localmente (tanto en el espacio como en el tiempo) en el continuo tiempo-espacio materia.
Se denomina marco en **co-movimiento** con la materia.

---

[203] Algunos físicos especialistas en la teoría cuántica de campos, como Zeldovich, relacionan la constante cosmológica con la energía de punto cero que debería aparecer en todas partes, inclusive en el espacio vacío. Sobre este tema puede consultarse la primera parte del libro editado por G. Fraser "*The new physics in the twenty first century*", Cambridge University Press, U.K, 2009.

Si bien el principio cosmológico débil así enunciado es compatible con algún tipo de evolución cósmica promedio, permanece el problema de cómo se conectan los tiempos y los espacios locales de todos esos marcos de referencia. El postulado de Weyl suministra una solución a este problema, basada en las características que parece presentar la expansión del Universo vista en 12.3.

Una vez que por promediado adecuado se eliminan los movimientos peculiares de las galaxias, el movimiento remanente parece ser un flujo expansivo isotrópico que puede denominarse **flujo de Hubble**.
El movimiento peculiar de las galaxias relativo al flujo de Hubble local presenta velocidades generalmente pequeñas[204] y prácticamente aleatorias.
La versión más conveniente del postulado de Weyl desde la perspectiva de los modelos cosmológicos, afirma que las líneas de universo de los observadores que se mueven con el flujo de Hubble, denominados **observadores fundamentales**, forman un haz regular de geodésicas temporaloides que no se intersecan en evento alguno, excepto, tal vez, en una singularidad situada en un pasado o en un futuro remotos.
El tiempo propio de cada observador fundamental se puede correlacionar con el tiempo propio de cualquier otro de tal modo que es posible sincronizarlos y definir un **tiempo cósmico universal** asignable a cada evento en el Universo.
Como consecuencia es posible identificar los eventos con el mismo tiempo cósmico, definiendo así en el continuo tiempo-espacio-materia unas variedades tridimensionales espacialoides que de acuerdo con el postulado cosmológico deben ser homogéneas e isótropas.
Cada una de estas hiper-superficies es el espacio físico correspondiente a un instante de tiempo cósmico (Figura 12.1).

[204] El fondo de radiación cósmica fósil de microondas presenta una ligera anisotropía (anisotropía dipolar) a partir de la cual se puede inferir el movimiento del sistema Solar, y la Tierra en particular, respecto del observador fundamental local. El Sol se mueve a 368±2 km/s y la velocidad de la Tierra en su órbita alrededor del Sol es la veinteava parte de esa velocidad. Los movimientos locales respecto de un observador fundamental son del orden de la milésima parte de la velocidad de la luz. La Vía Láctea, considerada en su totalidad, al parecer también se mueve respecto del observador fundamental local, aunque a muy baja velocidad respecto de la velocidad de la luz. Ver, por ejemplo, el libro de S. Serjeant "*Observational cosmology*", Cambridge University Press, U.K, 2010.

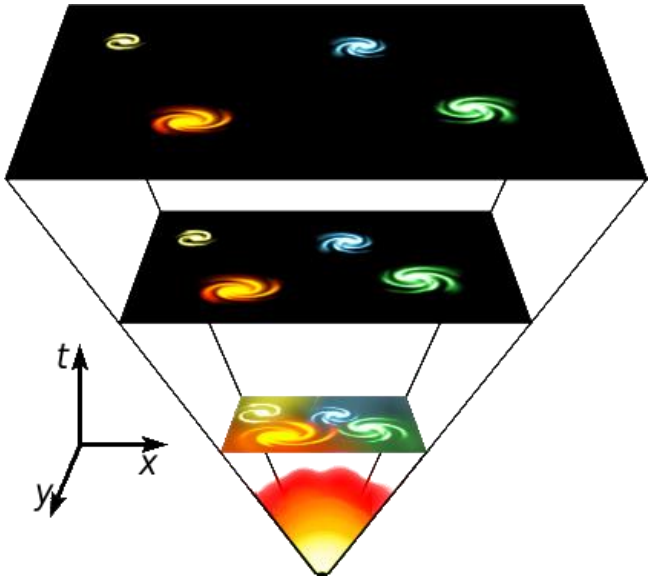


Figura 12.1: Representación del Flujo de Hubble mediante un tiempo-espacio-materia descomponible en variedades espacialoides (espacio físico) planas y un tiempo cósmico ortogonal al espacio físico. Cerca del vértice que indica el Big Bang se sitúa la fase inflacionaria seguida de la separación de la radiación y la materia.

Si el Universo evoluciona en gran escala a medida que transcurre el tiempo cósmico, entonces alguna propiedad geométrica de estas hiper-superficies espacialoides puede variar.
Veamos ahora cuál de ellas varía, cómo lo hace y qué significado posee esa variación.

12.4 Modelos basados en la métrica de Friedmann-Lemaître-Robertson-Walker

Del postulado de Weyl se desprende la posibilidad de seleccionar un tiempo universal o cósmico que denotaremos t tal que: $ds^2 = c^2 \cdot dt^2 - \sum_{i,k=1}^{3} g_{i,k}\left(t, x^1, x^2, x^3\right) \cdot dx^i \cdot dx^k$

A cada instante de tiempo cósmico le corresponde una variedad tridimensional (el "espacio") cuyo elemento de longitud depende de t y viene dado por la métrica:

$$dl^2 = \sum_{i,k=1}^{3} g_{i,k}\left(t, x^1, x^2, x^3\right) \cdot dx^i \cdot dx^k$$

Consideremos la variedad espacial en el instante $t = t_0$ y seleccionemos dos puntos cualesquiera $P(t_0)$ y $Q(t_0)$. En un Universo dinámico, en un instante t esos puntos se transforman en los puntos $P(t)$ y $Q(t)$ de la variedad espacial que representa al espacio físico en el nuevo instante de tiempo cósmico.
Suponiendo que estas variedades son homogéneas e isótropas, los cocientes de distancias $d(P(t), Q(t))$ entre pares de puntos correspondientes no deberían depender del tiempo cósmico, es decir $\dfrac{d(P(t), Q(t))}{d(P(t_0), Q(t_0))}$ debería ser constante para todo par de puntos.
Este requerimiento puede ser satisfecho si $g_{i,k}\left(t, x^1, x^2, x^3\right) = S^2(t) \cdot h_{i,k}\left(x^1, x^2, x^3\right)$

Aquí $S(t)$ es un cierto factor de escala que caracteriza la dinámica del Universo en el marco de este tipo de modelo matemático. Se halla a partir de las ecuaciones del campo gravitatorio como se verá posteriormente.
El intervalo tal como lo determinan los observadores fundamentales queda:

$$ds^2 = c^2 \cdot dt^2 - S^2(t) \cdot \sum_{i,k=1}^{3} h_{i,k}\left(x^1, x^2, x^3\right) \cdot dx^i \cdot dx^k$$

Ahora identificamos una variedad tridimensional estacionaria cuyo elemento de longitud viene dado por: $dl_0^2 = \sum_{i,k=1}^{3} h_{i,k}\left(x^1, x^2, x^3\right) \cdot dx^i \cdot dx^k$

Si ubicamos el origen de coordenadas en donde se halla un observador fundamental O, debido a la hipótesis de isotropía, la variedad presentará simetría esférica en torno al origen. Entonces se pueden introducir, con centro en O, una coordenada radial r y dos coordenadas angulares θ, φ de modo tal que: $dl_0^2 = g^2(r) \cdot dr^2 + r^2 \cdot \left(d\theta^2 + sen^2\theta \cdot d\varphi^2\right)$
Pero además de isótropa en O, la variedad es por hipótesis homogénea, y en consecuencia isótropa vista desde cualquier punto. Pero si todos sus puntos son puntos de isotropía, sabemos a partir de los resultados obtenidos en 9.5.4 para las variedades 3-D simétricas, que un único escalar K_0 caracteriza la curvatura en todo el espacio, y la función $g(r)$ debe ser igual a $g^2(r) = \dfrac{1}{1 - K_0 \cdot r^2}$

Si se introduce un radio de curvatura r_0 y un parámetro ε que puede tomar tres valores, -1, 0 y +1, la curvatura se puede re-escribir $K_0 = +\varepsilon \cdot \dfrac{1}{r_0^2}$

Entonces el intervalo se puede escribir finalmente así, en función del instante de tiempo cósmico y de las coordenadas r, θ, φ en co-movimiento[205] con la materia local:

$$ds^2 = c^2 \cdot dt^2 - S^2(t) \cdot \left(\left(1 - \varepsilon \cdot \frac{r^2}{r_0^2}\right)^{-1} \cdot dr^2 + r^2 \cdot \left(d\theta^2 + sen^2\theta \cdot d\varphi^2\right) \right)$$

Si se introduce $\tilde{r} = \dfrac{r}{r_0}$ en lugar de r y se define $r_0(t) = S(t) \cdot r_0$, la métrica queda:

$$ds^2 = c^2 \cdot dt^2 - r_o^2(t) \cdot \left(\left(\frac{1}{1 - \varepsilon \cdot \tilde{r}^2} \right) \cdot d\tilde{r}^2 + \tilde{r}^2 \cdot \left(d\theta^2 + sen^2\theta \cdot d\varphi^2\right) \right)$$

La coordenada r se toma con origen en cualquier punto del espacio físico, puesto que todos son equivalentes. La coordenada t es el tiempo propio en ese punto. La función no negativa $r_0(t)$ es una escala cósmica de longitud que permite representa la expansión o contracción del Universo. El parámetro ε puede tomar tres valores: -1 (espacio curvo abierto), 0 (espacio plano), y 1 (espacio curvo cerrado). La coordenada angular θ varía entre 0 y π, mientras que la coordenada angular φ varía entre 0 y 2π.
En estas coordenadas el tensor métrico posee las componentes: $g_{00} = c^2$

[205] Elegimos un marco de referencia para cada punto del espacio de tal modo que se mueva con la materia allí localizada: la velocidad de la materia en este **sistema de referencia intrínseco** es cero por definición.

$$g_{11} = -\frac{1}{1-\varepsilon\cdot\tilde{r}^2} \qquad g_{22} = -\tilde{r}^2 \qquad g_{33} = -\tilde{r}^2\cdot sen^2\theta \qquad g_{i,k} = 0 \quad i \neq k$$

Esta (pseudo)-métrica se denomina **métrica estándar**.

Una métrica equivalente[206] fue descubierta por Friedmann en 1922 y re-descubierta por Lemaître en 1927. Tanto Friedmann como Lemaître partieron de las ecuaciones de la teoría general de la relatividad para construir la métrica estándar y asumieron que no se podía desvincular la una de las otras. Pero en 1935, Howard Robertson y Arthur Walker demostraron que la métrica estándar es consecuencia de postular un continuo tiempo-espacio homogéneo e isótropo.

Cuando $\varepsilon = 0$, expresando el término relacionado con las coordenadas espaciales en su forma cartesiana, la métrica estándar se reduce a:

$$ds^2 = c^2\cdot dt^2 - S^2(t)\cdot\left(dx^2 + dy^2 + dz^2\right)$$

La presencia de un factor de escala que puede variar en el tiempo tiene como consecuencia que el tiempo-espacio tomado como un todo pueda estar curvado, aunque el espacio tridimensional asociado a las coordenadas de posición no lo esté.

Un factor de escala cambiante hace que el tiempo-espacio curvado correspondiente a la métrica estándar sea muy diferente[207] al espacio plano de Minkowski cuya métrica es, como sabemos: $ds^2 = c^2\cdot dt^2 - \left(dx^2 + dy^2 + dz^2\right)$

Para determinar la función $S(t)$ o su equivalente $r_0(t) = S(t)\cdot r_0(t_0)$, es necesario un modelo de gravitación. Antes de considerar cómo varía esta función con el tiempo, conviene fijar el instante y analizar algunos aspectos de las geometrías no euclidianas del espacio físico resultante.

---

[206] Ver, por ejemplo, el libro de H. Bondi "*Cosmology*", capítulo 10, Dover, New York, 2010.

[207] Con el fin de estudiar la curvatura del tiempo-espacio provocada por un factor de escala cambiante, consideraremos un caso simplificado de la métrica estándar, a saber: $ds^2 = c^2\cdot dt^2 - S^2(t)\cdot dx^2$ Poniendo $x^0 = c\cdot t$ $\quad x^1 = x$ $\quad g_{00} = 1$ $\quad g_{11} = -S^2\left(x^0\right)$ $\quad g_{ik} = 0$ (si $i \neq k$), entonces las componentes contravariantes del tensor métrico verifican $g^{00} = 1$ $\quad g^{11} = -\frac{1}{S^2\left(x^0\right)}$ $\quad g^{ik} = 0$ (si $i \neq k$). A partir de las componentes del tensor métrico se pueden hallar los coeficientes de conexión: $\Gamma^0{}_{00} = 0$ $\Gamma^0{}_{01} = \Gamma^0{}_{10} = 0$ $\quad \Gamma^0{}_{11} = -\frac{1}{2}\cdot g^{00}\cdot\frac{\partial g_{11}}{\partial x^0} = S\cdot\frac{dS}{dx^0}$ $\quad \Gamma^1{}_{00} = 0$ $\quad \Gamma^1{}_{01} = \Gamma^1{}_{10} = \frac{1}{2}\cdot g^{11}\cdot\frac{\partial g_{11}}{\partial x^0} = \frac{1}{S}\cdot\frac{dS}{dx^0}$ $\Gamma^1{}_{11} = 0$ A partir de los coeficientes de conexión, aplicando la fórmula para el tensor de curvatura de Riemann-Christoffel vista en 9.5.4, se obtienen las componentes $R^i{}_{jkl}$. En este caso la única componente no nula (a menos de simetrías) depende del tiempo y es: $R^0{}_{101}\left(x^0\right) = \frac{\partial\Gamma^0{}_{11}}{\partial x^0} - \Gamma^1{}_{10}\cdot\Gamma^0{}_{11} = S\left(x^0\right)\cdot\frac{d^2S\left(x^0\right)}{\left(dx^0\right)^2}$

12.4.1 Estudio estático de las geometrías a partir del término espacial del intervalo. Distancias propias y parámetro de Hubble.

Si fijamos un instante t y si introducimos el factor de escala $S(t)$ multiplicando la coordenada radial y continuamos llamando r a lo que ahora es $S(t)\cdot r$, el elemento de distancia espacial se puede escribir así, con $r_0(t)=S(t)\cdot r_0$:

$$dl^2 = \left(\left(1-\varepsilon\cdot\frac{r^2}{r_0{}^2(t)}\right)^{-1}\cdot dr^2 + r^2\cdot\left(d\theta^2 + sen^2\theta\cdot d\varphi^2\right)\right)$$

Consideremos la superficie de una esfera centrada en el origen y correspondiente a un valor constante de la coordenada r. El elemento de área es $r^2\cdot sen\theta\cdot d\theta\cdot d\varphi$, de modo que su área es igual a $4\pi\cdot r^2$.

En el caso de una circunferencia con centro en el origen de coordenadas (caracterizada por un valor de r constante) y situada en el plano $\varphi=0$, el elemento de longitud es $dl = r\cdot d\theta$ de modo que su longitud es igual a $2\pi\cdot r$.

No obstante el radio $\tilde{\rho}$ de estas figuras no es igual a r sino a $\displaystyle\int_0^r \frac{d\xi}{\sqrt{1-\varepsilon\cdot\frac{\xi^2}{r_0^2(t)}}}$

Se ve que si $\varepsilon=1$ (espacio curvo cerrado), entonces $\tilde{\rho}$ es mayor que r; si $\varepsilon=0$ (espacio plano) entonces $\tilde{\rho}$ es igual que r; y si $\varepsilon=-1$ (espacio curvo abierto) entonces $\tilde{\rho}$ es menor que r.

Estas variedades espaciales 3-D son simétricas respecto del origen de coordenadas, del tipo que estudiamos en detalle en el capítulo 9 (en las sub-secciones 9.4.6, 9.5.3 y 9.5.4), con un elemento de línea del tipo $dl^2 = g^2(r)\cdot dr^2 + r^2\cdot\left(d\theta^2 + sen^2\theta\cdot d\varphi^2\right)$

Como en este caso se trata de un espacio isótropo en todas partes: $g(r)=\dfrac{1}{\sqrt{1-\varepsilon\cdot\frac{r^2}{r_0{}^2}}}$

Entonces la curvatura de la variedad 3-D, que representa al espacio físico en este modelo, se caracteriza por el mismo escalar en todos los puntos considerados en el mismo instante de tiempo universal, pero esa curvatura uniforme del espacio físico puede variar con el instante de acuerdo con la fórmula: $K(t)=\dfrac{\varepsilon}{r_0^2(t)}$

Espacio esférico

Consideremos el caso para el cual $\varepsilon=1$. En ese caso la distancia radial al origen $\tilde{\rho}$ viene dada por la fórmula:

$$\tilde{\rho} = \int_0^r \frac{d\xi}{\sqrt{1-\frac{\xi^2}{r_0^2(t)}}} = r_0(t)\cdot arcsen\left(\frac{r}{r_0(t)}\right)$$

Poniendo $\dfrac{r}{r_0} = sen\chi$ el elemento de longitud y la distancia al origen vienen dados en función de χ: $dl^2 = r_0^2\cdot\left(d\chi^2 + sen^2\chi\cdot\left(sen^2\theta\cdot d\varphi^2 + d\theta^2\right)\right)$ $\tilde{\rho} = r_0\cdot\chi$

Como $\frac{r}{r_0}$ debe estar comprendido entre 0 y 1, el ángulo χ debe variar entre 0 y 2π .

En términos de la coordenada angular χ el área de una esfera de radio $\tilde{\rho} = r_0 \cdot \chi$ es $S_e = 4\pi \cdot r^2 = 4\pi \cdot r_0^2 \cdot sen^2 \chi$ y la longitud de una circunferencia como la considerada previamente es $l_c = 2\pi \cdot r = 2\pi \cdot r_0 \cdot sen\chi$.

Como $sen\chi = \chi - \frac{1}{6} \cdot \chi^3 + o(\chi^3)$, donde $o(\chi^3)$ tiende a cero más rápido que χ^3, la longitud de la circunferencia se puede expresar así, con $\varepsilon_c = 2\pi . r_0 \cdot o(\chi^3)$:

$$l_c = 2\pi \cdot r_0 \cdot \left(\chi - \frac{1}{6} \cdot \chi^3 + o(\chi^3) \right) = 2\pi \cdot r_0 \cdot \chi - \frac{\pi}{3} \cdot \frac{1}{r_0^2} \cdot (r_0^3 \cdot \chi^3) + \varepsilon_c$$

Entonces, en términos del radio $\tilde{\rho} = r_0 \cdot \chi$, obtenemos finalmente la fórmula: $l_c = 2\pi \cdot \tilde{\rho} - \frac{\pi}{3} \cdot \frac{1}{r_0^2(t)} \cdot \tilde{\rho}^3 + \varepsilon_c$ Comparando con la fórmula que relaciona la longitud de la circunferencia con su radio y con su curvatura de Gauss, vista en 9.1.2, $l_c = 2\pi \cdot \tilde{\rho} - \frac{\pi}{3} \cdot K \cdot \tilde{\rho}^3 + \varepsilon_c$ resulta que para $\varepsilon = 1$ el espacio posee en el punto considerado la curvatura de Gauss positiva $K(t) = \frac{1}{r_0^2(t)}$.

El área de una superficie esférica $S_e = 4\pi \cdot r^2 = 4\pi \cdot r_0^2 \cdot sen^2 \chi$, de radio $\tilde{\rho} = r_0 \cdot \chi$, crece desde 0 cuando $\chi = 0$ ($\tilde{\rho} = 0$) hasta un máximo igual a $S_e = 4\pi \cdot r_0^2$ para $\chi = \frac{\pi}{2}$ ($\tilde{\rho} = r_0 \cdot \frac{\pi}{2}$) y luego decrece hasta anularse nuevamente, ahora para $\chi = \pi$ ($\tilde{\rho} = r_0 \cdot \pi$).

Este comportamiento de las superficies de las esferas con centro en un mismo punto del espacio, en tres dimensiones, es análogo al comportamiento de las circunferencias contenidas en una superficie esférica (que es un espacio bidimensional curvado) y centradas en un mismo punto, a medida de que el radio aumenta, medido por las longitudes de los arcos de círculo máximo (es decir de las geodésica en la esfera) que va del centro hasta la circunferencia. La longitud primero aumenta hasta alcanzar un máximo en el ecuador y luego disminuye hasta anularse al alcanzar el polo opuesto al puntos tomado como centro.

Si introducimos un espacio euclidiano de 4 dimensiones de coordenadas cartesianas ortogonales x_1, x_2, x_3, x_4, el espacio físico que corresponde al caso $\varepsilon = 1$ en la métrica de Friedmann-Lemaître-Robertson-Walker puede interpretarse como una **cáscara esférica de 3 dimensiones**, de radio r_0 , centrada en el origen de ese espacio y dada por la ecuación $x_1^2 + x_2^2 + x_3^2 + x_4^2 = r_0^2$. Como vimos en 9.4.5, introduciendo un sistema de coordenadas esféricas la cáscara de radio r_0 se puede representar así (χ entre 0 y 2π; θ entre 0 y π ; φ entre 0 y 2π):

$x_1 = r_0 \cdot sen\chi \cdot sen\theta \cdot \cos\varphi$ $\quad$ $x_2 = r_0 \cdot sen\chi \cdot sen\theta \cdot sen\varphi$ $\quad$ $x_3 = r_0 \cdot sen\chi \cdot \cos\theta$

$x_4 = r_0 \cdot \cos\chi$

Por tanto es una variedad finita, cerrada y en sí misma ilimitada (recorriéndola por adentro no se encuentra una frontera).

Partiendo de un punto cualquiera (tomado como polo), el punto más alejado de la variedad respecto de este polo se encuentra a la distancia $\tilde{\rho} = r_0 \cdot \pi$, en el polo opuesto.
El volumen de esa cáscara, o sea, **el volumen del cosmos** según este modelo de Universo cerrado, es $V_U(t) = r_0^3(t) \cdot \int_0^{\pi}\int_0^{\pi}\int_0^{2\pi} sen^2\chi \cdot sen\theta \cdot d\varphi \cdot d\theta \cdot d\chi = 2\pi^2 \cdot r_0^3(t) \cdot$

El volumen de una región de radio $\tilde{\rho} = r_0 \cdot \chi_0$ es $V(\chi_0) = 2\pi \cdot r_0^3\left(\chi_0 - \frac{1}{2} \cdot sen(2 \cdot \chi_0)\right)$

Espacio hiperbólico
Cuando $\varepsilon = -1$, la distancia radial al origen viene dada por la fórmula:

$$\tilde{\rho} = \int_0^r \frac{d\xi}{\sqrt{1 + \frac{\xi^2}{r_0^2(t)}}} = r_0(t) \cdot arcsenh\left(\frac{r}{r_0(t)}\right)$$

Introduciendo $\frac{r}{r_0} = senh\chi$, se obtiene nuevamente $\tilde{\rho} = r_0 \cdot \chi$. Pero ahora χ puede variar entre 0 y $+\infty$. Es un universo abierto.
El área de una superficie esférica de radio $\tilde{\rho} = r_0 \cdot \chi$ es $S_e = 4\pi \cdot r^2 = 4\pi \cdot r_0^2 \cdot senh^2\chi$, y tiende a infinito cuando χ crece. El volumen de este Universo es infinito.
Como $senh\chi = \chi + \frac{1}{6} \cdot \chi^3 + o(\chi^3)$, donde $o(\chi^3)$ tiende a cero más rápido que χ^3, la longitud de la circunferencia $l_c = 2\pi \cdot r = 2\pi \cdot r_0 \cdot senh\chi$ se puede expresar así, con $\varepsilon_c = 2\pi . r_0 \cdot o(\chi^3)$: $l_c = 2\pi \cdot r_0 \cdot \left(\chi + \frac{1}{6} \cdot \chi^3 + o(\chi^3)\right) = 2\pi \cdot r_0 \cdot \chi + \frac{\pi}{3} \cdot \frac{1}{r_0^2} \cdot (r_0^3 \cdot \chi^3) + \varepsilon_c$
Comparando nuevamente con la fórmula que relaciona la longitud de la circunferencia con su radio y con su curvatura de Gauss, $l_c = 2\pi \cdot \tilde{\rho} - \frac{\pi}{3} \cdot K \cdot \tilde{\rho}^3 + \varepsilon_c$ resulta que para $\varepsilon = -1$ el espacio posee en el punto considerado la curvatura de Gauss negativa $K(t) = -\frac{1}{r_0^2(t)}$.
Si introducimos un espacio euclidiano de 4 dimensiones de coordenadas cartesianas ortogonales x_1, x_2, x_3, x_4, el espacio físico que corresponde al caso $\varepsilon = -1$ en la métrica de Friedmann-Lemaître-Robertson-Walker puede interpretarse, como vimos en 9.4.5, como una **variedad hiperbólica de 3 dimensiones**, dada por la ecuación $x_1^2 + x_2^2 + x_3^2 - x_4^2 = r_0^2$.
Introduciendo un nuevo sistema de coordenadas χ, θ, φ se puede representar así, teniendo en cuenta que ahora χ varía entre 0 y $+\infty$, mientras que las coordenadas angulares varían como de costumbre, es decir θ entre 0 y π; φ entre 0 y 2π:
$x_1 = r_0 \cdot \cosh\chi \cdot sen\theta \cdot \cos\varphi$ $\quad x_2 = r_0 \cdot \cosh\chi \cdot sen\theta \cdot sen\varphi$ $\quad x_3 = r_0 \cdot \cosh\chi \cdot \cos\theta$
$x_4 = r_0 \cdot senh\chi$
Puesto que el coseno hiperbólico y el seno hiperbólico son funciones no acotadas (tienden a $+\infty$ cuando χ tiende a $+\infty$), tenemos un espacio que, además de **uniformemente curvado e ilimitado, es un espacio abierto** porque nos podemos alejar tanto como queramos a partir de un punto, puesto que a diferencia de lo que

ocurre en el caso con $\varepsilon = +1$, ahora no hay una cota superior para las distancias entre puntos del espacio: el volumen del espacio es infinito[208].
Si la coordenada radial está acotada $r \leq r_F$, el volumen de la región espacial así definida se puede calcular mediante la fórmula:

$$V_{r_F}(t) = r_0^3(t) \cdot \int_0^{r_F}\int_0^{\pi}\int_0^{2\pi} senh^2\chi \cdot sen\theta \cdot d\varphi \cdot d\theta \cdot d\chi = 2\pi^2 \cdot r_0^3(t) \cdot \left(\frac{r_F(t)}{r_0(t)} \cdot \sqrt{1 + \frac{r_F(t)^2}{r_0(t)^2}} - arcsenh \frac{r_F(t)}{r_0(t)} \right)$$

Si se supone una expansión homogénea cabe esperar que $\frac{r_F(t)}{r_0(t)}$ permanezca constante.

Espacio plano
Cuando $\varepsilon = 0$, $\tilde{\rho} = r$ y se tiene el espacio euclidiano tridimensional, es decir el espacio ordinario en el que organizamos nuestras percepciones.

Parámetro de Hubble
La distancia entre dos observadores en co-movimiento local con la materia, cuando uno de ellos se encuentra situado en el origen, para el mismo instante de tiempo cósmico viene dada por $\tilde{\rho}$. Se mide a lo largo de la geodésica que parte del origen en una dirección especificada por los ángulos θ y φ.

Cuando $\frac{r}{r_0(t)}$ se mantiene constante, de la fórmula $\tilde{\rho}(t) = r_0(t) \cdot \int_0^{\frac{r}{r_0(t)}} \frac{d\zeta}{\sqrt{1 - \varepsilon \cdot \zeta^2}}$ se desprende que la distancia $\tilde{\rho}$ entre los observadores aumenta con una velocidad $\frac{d}{dt}\tilde{\rho} = \left(\frac{1}{r_0(t)} \cdot \frac{dr_0(t)}{dt} \right) \cdot \tilde{\rho}(t) = H(t) \cdot \tilde{\rho}(t)$ donde $H(t) = \frac{1}{r_0(t)} \cdot \frac{dr_0(t)}{dt}$ se conoce como **parámetro de Hubble**. La velocidad radial de separación de dos observadores en co-movimiento local con la materia resulta proporcional a su distancia y a ese parámetro característico de la expansión del Universo, expresado en términos de la velocidad relativa de variación del radio de curvatura del modelo de Universo. Se aplica a los tres casos posibles: $\varepsilon = 1$ (espacio curvo cerrado), $\varepsilon = 0$ (espacio plano) y si $\varepsilon = -1$ (espacio curvo abierto).

12.4.2 Estudio a partir de la métrica completa. El modelo estándar Λ-CDM

Para hallar la función $r_0(t)$ que determina la curvatura del modelo de Universo se sustituyen las componentes del tensor métrico correspondiente a la métrica estándar (métrica de Friedmann- Lemaître-Robertson-Walker) en las ecuaciones del campo gravitatorio. Pero antes de hacer esto es conveniente efectuar el cambio de coordenadas espaciales visto en 9.4.6 y también al final de 9.5.4, que permite una representación conforme de la variedad espacial en E_3, poniendo $\tilde{r} = \frac{r_1}{1 + \varepsilon \cdot \frac{r_1^2}{4}}$:

[208] Como se mostró en 9.4.5, si se elimina una de las coordenadas espaciales haciendo $x_3 \equiv 0$, la ecuación de la variedad se reduce a $x_1^2 + x_2^2 - x_4^2 = r_0^2$ que en un espacio euclidiano tridimensional corresponde a la superficie conocida como hiperboloide de una hoja.

$$dl_1^2 = A^2 \cdot \left(dr_1^2 + r_1^2 \cdot d\theta^2 + r_1^2 \cdot sen^2\theta \cdot d\varphi^2\right) \qquad A^2(r_1) = \frac{1}{1+\varepsilon \cdot \frac{r_1^2}{4}}$$

La métrica de Friedmann-Lemaître-Robertson-Walker se lleva así a la **forma canónica**:

$$ds^2 = c^2 \cdot dt^2 - r_o^2(t) \cdot A^2(r_1) \cdot \left(dr_1^2 + r_1^2 \cdot \left(d\theta^2 + sen^2\theta \cdot d\varphi^2\right)\right)$$

Aquí $r_0(t)$ es un factor global de escala cósmica, $A^2(r_1) = \frac{1}{1+\varepsilon \cdot \frac{r_1^2}{4}}$ es un factor local de curvatura y $dr_1^2 + r_1^2 \cdot \left(d\theta^2 + sen^2\theta \cdot d\varphi^2\right)$ es una métrica euclidiana en el espacio tangente en el punto considerado de la variedad tetra-dimensional definida por la métrica. La única función desconocida es la dependencia del tiempo del factor de escala cósmico.

Cuando r_1 se aproxima a cero, $A^2(r_1)$ se aproxima a 1, y r_1 se aproxima a r.

El elemento de línea correspondiente a la métrica de Friedmann-Lemaître-Robertson-Walker se aproxima al elemento de línea:

$$ds^2 = c^2 \cdot dt^2 - \left(dr^2 + r^2 \cdot d\theta^2 + r^2 \cdot sen^2\theta \cdot d\varphi^2\right)$$

Este último define un espacio de Minkowski (marco de referencia inercial) tangente en el origen a la variedad curvada.

Para hallar la función $r_0(t)$ que determina la curvatura del modelo de Universo se sustituyen las componentes del tensor métrico correspondiente a la métrica estándar puesta en su forma canónica, en las ecuaciones del campo gravitatorio:

$$G^{i,k} = R^{i,k} - \frac{1}{2} \cdot R \cdot g^{i,k} = -\kappa \cdot T^{i,k} \quad (i,k = 0,1,2,3) \qquad \kappa = \frac{8\pi}{c^4} \cdot G$$

Como se supone que el observador se mueve localmente junto con el fluido cósmico homogéneo e isótropo, caracterizado por una densidad propia y una presión uniformes $\rho(t)$ y $p(t)$ respectivamente, las componentes del tensor de energía impulsión pueden escribirse así: $T^{0,0} = \rho(t) \cdot c^2 \qquad T^{1,1} = T^{2,2} = T^{3,3} = p(t) \qquad T^{i,k} = 0$ si $i \neq k$

Sustituyendo en las ecuaciones para el campo gravitatorio esta versión del tensor de energía impulsión junto con las componentes del tensor métrico correspondiente a la métrica de Friedmann-Lemaître-Robertson-Walker, después de numerosos cálculos se obtienen las siguientes ecuaciones que relacionan el factor de escala con la presión p y la densidad ρ del Universo considerado como un fluido homogéneo e isótropo[209]:

$$\left(\frac{1}{r_0(t)} \cdot \frac{dr_0(t)}{dt}\right)^2 = \frac{8\pi \cdot G}{3} \cdot \rho(t) - \frac{\varepsilon \cdot c^2}{r_0^2(t)} \qquad \frac{1}{r_0(t)} \cdot \frac{d^2 r_0(t)}{dt^2} = -\frac{4\pi \cdot G}{3} \cdot \left(\rho(t) + \frac{3 \cdot p(t)}{c^2}\right)$$

Estas ecuaciones pueden denominarse **primera** y **segunda ecuación de Friedmann-Lemaître**.

---

[209] E. Harrison "*Cosmology: the science of the Universe*", Cambridge University Press, Londres, 2008. L. Landau y E. Lifchitz "*Teoría del campo*", Reverté, Barcelona, 1982.

La interpretación actual de los parámetros del tensor de energía-impulsión (ρ y p) es en buena medida diferente a la interpretación que se les daba en los primeros modelos cosmológicos de Einstein, de Sitter, Friedmann y Lemaître, sobre todo debido a la inclusión de la materia oscura y de la energía oscura.
En este capítulo revisaremos una familia de modelos conocidos como modelos de Friedmann-Lemaître-Robertson-Walker. A diferencia de lo que se hizo cuando se construyeron los primeros modelos cosmológicos, en la formulación de los modelos modernos se tiene en cuenta una mezcla de fluidos cósmicos, cada uno con su ecuación de estado. Con estos modelos más sofisticados se obtienen resultados al parecer mejores en relación con la posible edad del universo y se pueden describir detalles de la dinámica de expansión que quedan fuera del alcance de los modelos antiguos.

El fluido cósmico se modela como una mezcla aditiva de tres fluidos ideales componentes: el formado por la materia (incluyendo la materia ordinaria y la materia negra) de densidad $\rho_m(t)$ y presión $p_m(t)$, el formado por la radiación de densidad $\rho_r(t)$ y presión $p_r(t)$, y el formado por la energía negra de densidad ρ_Λ y presión p_Λ.
Entonces: $\rho(t) = \rho_m(t) + \rho_r(t) + \rho_\Lambda$ y $p(t) = p_m(t) + p_r(t) + p_\Lambda$

Para cada uno de los tres fluidos se introduce una ecuación de estado del tipo que corresponde a un fluido perfecto, $p = w \cdot \rho \cdot c^2$ donde p es la presión, ρ es la densidad tal como se definen para un marco de referencia que se mueve localmente con el fluido y w es un parámetro a-dimensionado.
Para la materia $w = w_m$ toma el valor 0 si se considera que se comporta como un polvo, o de lo contrario adopta un valor positivo. En lo que sigue asumiremos que $w_m = 0$.
Para la radiación $w = w_r = \frac{1}{3}$. Para la energía oscura[210] $w = w_\Lambda = -1$.

Como consecuencia de estas suposiciones:

$$p(t) = p_m(t) + p_r(t) + p_\Lambda = 0 + \frac{1}{3} \cdot \rho_r(t) \cdot c^2 - \rho_\Lambda \cdot c^2$$

Teniendo en cuenta esta fórmula que expresa la presión en términos de las densidades, la segunda ecuación de Friedmann-Lemaître, también denominada ecuación de la aceleración, se reduce a la siguiente:

$$\frac{1}{r_0(t)} \cdot \frac{d^2 r_0(t)}{dt^2} = -\frac{4\pi \cdot G}{3} \cdot \left(\rho_m(t) + 2 \cdot \rho_r(t) - 2 \cdot \rho_\Lambda\right)$$

[210] A veces se supone que la energía oscura se manifiesta como un fluido cuya ecuación de estado adopta la forma $p_Q = w_Q \cdot \rho_Q \cdot c^2$ con $w_Q < -\frac{1}{3}$. Una vez que la dinámica del cosmos en expansión sale de las etapas en las que domina la radiación (primero) y domina la materia (después) el Universo entraría en una etapa cuya dinámica es dominada por la energía oscura, acelerando su expansión si en la ecuación de estado de la materia oscura se verifica $w_Q < -\frac{1}{3}$. A un fluido como éste algunos cosmólogos se complacen en denominarlo **quintaesencia**, en recuerdo del quinto elemento del antiguo modelo cosmológico, que se suponía era la materia de la que se encontraban hechos los astros y no tenía equivalente en el mundo sublunar que habitamos.

Como la densidad de energía oscura se considera positiva, de la segunda ecuación de Friedmann-Lemaître se desprende que esta forma de energía tiende a acelerar una expansión del Universo, mientras que el efecto de las densidades positivas tanto de la materia como de la radiación es desacelerarla.
La denominada materia oscura fría, a diferencia de la energía oscura, se supone que se comporta como la materia ordinaria desde el punto de vista de las interacciones gravitatorias y tiende a frenar una expansión del Universo.

La primera ecuación de Friedmann-Lemaître puede escribirse en términos de la suma de las densidades de los tres fluidos: $\left(\frac{1}{r_0(t)}\cdot\frac{dr_0(t)}{dt}\right)^2=\frac{8\pi\cdot G}{3}\cdot\left(\rho_m(t)+\rho_r(t)+\rho_\Lambda\right)-\frac{\varepsilon\cdot c^2}{r_0^2(t)}$

Como $\frac{1}{r_0(t)}\cdot\frac{dr_0(t)}{dt}=H(t)$ es el parámetro de Hubble introducido al final de 12.4.1, esta ecuación se puede reformular así: $\frac{\varepsilon\cdot c^2}{r_0^2(t)}=\frac{8\pi\cdot G}{3}\cdot\left(\rho(t)-\frac{3\cdot H^2(t)}{8\pi\cdot G}\right)$

Definiendo una **densidad crítica** $\rho_c(t)=\frac{3\cdot H^2(t)}{8\pi\cdot G}$ resulta que el valor de ε y por tanto, de acuerdo con esta clase de modelos, la geometría del Universo, depende de si la densidad total del fluido cósmico es mayor, igual o menor que ese valor crítico.
Se introducen los denominados parámetros de densidad para la materia, la radiación y la energía oscura: $\Omega_m(t)=\frac{\rho_m(t)}{\rho_c(t)}$ $\Omega_r(t)=\frac{\rho_r(t)}{\rho_c(t)}$ $\Omega_\Lambda(t)=\frac{\rho_\Lambda}{\rho_c(t)}$

Como $\frac{\varepsilon\cdot c^2}{r_0^2(t)\cdot H^2(t)}=\Omega_m(t)+\Omega_r(t)+\Omega_\Lambda(t)-1$, resulta finalmente lo siguiente, denominando **espacio físico** a las **hiper-superficies espacialoides** del tiempo-espacio-materia:

Si $\Omega_m(t)+\Omega_r(t)+\Omega_\Lambda(t)>1$ el espacio físico es finito (cerrado) y con curvatura positiva (caso esférico, $\varepsilon=+1$).

Si $\Omega_m(t)+\Omega_r(t)+\Omega_\Lambda(t)=1$ el espacio físico es infinito (abierto) y con curvatura nula (caso plano, $\varepsilon=0$).

Si $\Omega_m(t)+\Omega_r(t)+\Omega_\Lambda(t)<1$ el espacio físico es infinito (abierto) y con curvatura negativa (caso hiperbólico, $\varepsilon=-1$).

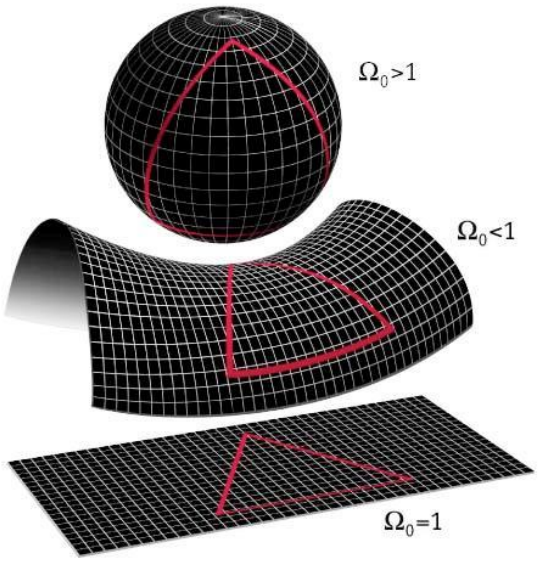


Figura 12.2 Tres geometrías posibles para el espacio físico según el valor del parámetro $\Omega_0(t)=\frac{\rho_m+\rho_r+\rho_\Lambda}{\rho_c}$. Para representarlas se utilizan variedades bidimensionales inmersas en E3.
En cada una de esas variedades se indica un triángulo formado por geodésicas, cuya suma de ángulos se relaciona con la curvatura local y con el área del triángulo como se mostró en 9.1.2.

Es posible relacionar las densidades de materia y de radiación en un instante de tiempo cósmico con el radio de curvatura del Universo $r_0(t)$. Consideremos un volumen propio muy grande de espacio físico en un instante de tiempo cósmico $t=t_0$, cuando las densidades de materia y de radiación son $\rho_m(t_0)$ y $\rho_r(t_0)$. El radio del Universo en ese instante inicial es $r_0(t_0)$.
En otro instante de tiempo cósmico el radio será $r_0(t)$ y el volumen propio considerado habrá variado respecto del inicial en un factor $\left(\frac{r_0(t)}{r_0(t_0)}\right)^3$.

Si, como sugirió O. Heckmann[211], la materia y la radiación no se inter-convierten en cantidades significativas, el volumen modificado debe contener en su interior el mismo número de partículas de materia, cada una con su propia masa no afectada por la expansión o contracción uniforme. Entonces cabe esperar que la relación entre las densidades $\rho_m(t)$ y $\rho_m(t_0)$ sea: $\rho_m(t)=\rho_m(t_0)\cdot\left(\frac{r_0(t_0)}{r_0(t)}\right)^3$

En el caso de la radiación que ocupa el volumen considerado, su densidad inicial $\rho_r(t_0)$ se modifica en un factor $\left(\frac{r_0(t)}{r_0(t_0)}\right)^4$ debido a un cambio en el radio de curvatura del Universo, por el siguiente motivo. La densidad de la radiación se define a partir de su densidad de energía $e_r(t_0)$ como $\rho_r(t_0)=\frac{e_r(t_0)}{c^2}$, puesto que, a diferencia de las

[211] O. Heckmann "*Theorien der Kosmologie*", Springer-Verlag, Berlín, 1942.

partículas de materia, su densidad de reposo es nula. La densidad de energía de la radiación se compone de las contribuciones de los cuantos cuya energía es $h \cdot f = \frac{h \cdot c}{\lambda}$.

Al variar las dimensiones del volumen de una porción muy grande del Universo, varía la longitud de onda λ de modo que $\lambda(t) = \lambda(t_0) \cdot \left(\frac{r_0(t)}{r_0(t_0)} \right)$.

Entonces la energía de los cuantos varía como $\left(\frac{r_0(t_0)}{r_0(t)} \right)$ desde el instante de tiempo cósmico t_0 al instante t.

El número de corpúsculo de radiación por unidad de volumen varía como $\left(\frac{r_0(t)}{r_0(t_0)} \right)^3$.

Combinando el efecto sobre la densidad con el efecto sobre la energía de los corpúsculos se obtiene: $\rho_r(t) = \rho_r(t_0) \cdot \left(\frac{r_0(t_0)}{r_0(t)} \right)^4$

De las dependencias entre las densidades de materia y de radiación y el factor de escala cósmico, en ausencia de inter-conversión entre materia y radiación y de transferencia de calor entre las regiones consideradas, se deduce que[212]: $\frac{\rho_r(t)}{\rho_r(t_0)} \cdot \left(\frac{\rho_m(t)}{\rho_m(t_0)} \right)^{-\frac{4}{3}} = 1$

Teniendo todo esto en cuenta, se puede plantear una versión simplificada de las ecuaciones de Friedmann- Lemaître:

$$\left(\frac{1}{r_0(t)} \cdot \frac{dr_0(t)}{dt} \right)^2 = \frac{8\pi \cdot G}{3} \cdot \left[\rho_m(t_0) \cdot \left(\frac{r_0(t_0)}{r_0(t)} \right)^3 + \rho_r(t_0) \cdot \left(\frac{r_0(t_0)}{r_0(t)} \right)^4 + \rho_\Lambda \right] - \frac{\varepsilon \cdot c^2}{r_0^2(t)}$$

$$\frac{1}{r_0(t)} \cdot \frac{d^2 r_0(t)}{dt^2} = -\frac{4\pi \cdot G}{3} \cdot \left[\rho_m(t_0) \cdot \left(\frac{r_0(t_0)}{r_0(t)} \right)^3 + 2 \cdot \rho_r(t_0) \cdot \left(\frac{r_0(t_0)}{r_0(t)} \right)^4 - 2 \cdot \rho_\Lambda \right]$$

Esta forma simplificada de las ecuaciones de Friedmann-Lemaître permite una descripción analítica de algunos casos límite de la dinámica global del Universo[213].

Modelo plano con radiación dominante.

Supongamos que $\varepsilon = 0$ [214] y que las densidades de materia, de radiación y de energía oscura en el instante de referencia $t = t_0$ no se anulan.

---

[212] R. Alpher y R. Herman "Evolution of the Universe", Nature, volumen 162, 1948, pp. 774-775.

[213] Una descripción más completa de las posibles soluciones de la familia de modelos de Friedmann-Lemaître-Robertson-Walker requiere completar los detalles de cada modelo específico y emplear métodos numéricos en un contexto de simulación digital dinámica.

Cuando $\frac{r_0(t)}{r_0(t_0)}$ es lo bastante pequeño, la primera ecuación de Friedmann-Lemaître se puede aproximar por ésta: $\frac{1}{r_0(t)}\cdot\frac{dr_0(t)}{dt}\approx\sqrt{\frac{8\pi\cdot G}{3}\cdot\rho_r(t_0)}\cdot\left(\frac{r_0(t_0)}{r_0(t)}\right)^2$

La solución es de la forma: $r_0(t)\approx r_0(t_0)\cdot\sqrt{2\cdot H(t_0)\cdot t}$ con $H(t_0)=\sqrt{\frac{8\pi\cdot G}{3}\cdot\rho_r(t_0)}$

Entonces: $\frac{1}{r_0(t)}\cdot\frac{dr_0(t)}{dt}=H(t)=\frac{1}{2\cdot t}$

Se suele aceptar que, luego de una fase de expansión extraordinariamente rápida, denominada inflación, que se considera más adelante en este capítulo (en la sección 12.5), la dinámica del Universo estuvo dominada por la radiación, hasta que pasó a estar dominado por la materia y sobre todo por la energía oscura.

Modelo plano con materia y energía oscura dominantes: el modelo estándar Λ-CDM.

Cuando $\frac{r_0(t)}{r_0(t_0)}$ es lo bastante grande como para que el término de radiación pueda ser despreciado, la primera ecuación de Friedmann-Lemaître con $\varepsilon=0$ se puede aproximar por ésta: $\left(\frac{1}{r_0(t)}\cdot\frac{dr_0(t)}{dt}\right)\approx\sqrt{\frac{8\pi\cdot G}{3}\cdot\left[\rho_m(t_0)\cdot\left(\frac{r_0(t_0)}{r_0(t)}\right)^3+\rho_\Lambda\right]}$

Si el término asociado con la energía oscura pudiera despreciarse frente al término correspondiente a la materia, se obtiene un modelo como el propuesto en 1932 por Einstein y por de Sitter: $r_0(t)\approx r_0(t_0)\cdot\left(\frac{3}{2}\cdot t\right)^{\frac{2}{3}}$ $H(t_0)=\sqrt{\frac{8\pi\cdot G}{3}\cdot\rho_m(t_0)}$

$$\frac{1}{r_0(t)}\cdot\frac{dr_0(t)}{dt}=H(t)=\frac{2}{3\cdot t}$$

Este modelo describe una segunda etapa de evolución del Universo, con una dinámica dominada por la materia. La evidencia acumulada a partir de 1980 condujo a su abandono, a fines del siglo XX, y su sustitución por modelos que daban prioridad a los efectos debidos a la energía oscura.

Cuando $\frac{r_0(t)}{r_0(t_0)}$ crece lo suficiente, se pueden despreciar los términos asociados tanto con la radiación como con la materia.

En ese caso la primera ecuación de Friedmann-Lemaître con $\varepsilon=0$ se puede aproximar por: $\left(\frac{1}{r_0(t)}\cdot\frac{dr_0(t)}{dt}\right)\approx\sqrt{\frac{8\pi\cdot G}{3}\cdot\rho_\Lambda}$

---

[214] Aunque las hipersuperficies espacialoides que forman el espacio físico posean curvatura nula, como en este caso en el cual $\varepsilon=0$, el tiempo-espacio-materia tetra-dimensional está curvado debido al factor $r_0^2(t)$ que aparece en la métrica de Friedmann-Lemaître-Robertson-Walker.

La solución es una exponencial: $r_0(t) \approx r_0(t_0) \cdot \exp[H(t_0) \cdot (t - t_0)]$ donde $H(t) = H(t_0) = \sqrt{\frac{8\pi \cdot G}{3} \cdot \rho_\Lambda}$

Este modelo no resulta adecuado para describir la etapa actual del Universo, pero podría describir un futuro lejano, si la expansión continuara hasta reducir las densidades de materia y radiación a niveles despreciables en comparación con la densidad de energía oscura. Es matemáticamente idéntico al modelo propuesto por de Sitter en 1917, aunque su significado físico es diferente.

También ha sido utilizado para describir la etapa inflacionaria, extremadamente breve, una vez surgido el Universo a partir de lo que se suele denominar una singularidad en el modelo del Big Bang.

El modelo matemático más simple de la dinámica cósmica, que cuenta con la mayor aceptación en el momento actual, es el **modelo plano** que involucra a la vez el aporte de la **materia** (**ordinaria** y **oscura**) y de la **energía oscura**. Como la materia oscura (cold dark matter) y la energía oscura (a través del parámetro Λ) dominan la cinética, se lo conoce como modelo Λ-CDM. Para deducir el comportamiento temporal del factor de escala se integra por separación de variables la ecuación del modelo plano con materia oscura y energía oscura dominantes:

$$\left(\frac{1}{r_0(t)} \cdot \frac{dr_0(t)}{dt}\right) \approx \sqrt{\frac{8\pi \cdot G}{3} \cdot \left[\rho_m(t_0) \cdot \left(\frac{r_0(t_0)}{r_0(t)}\right)^3 + \rho_\Lambda\right]}$$

Se supone que en el instante inicial $r_0(0) = 0$. Se obtiene así la siguiente fórmula para el factor de escala del Universo: $\frac{r_0(t)}{r_0(t_0)} = S(t) = \left(\frac{\rho_m(t_0)}{\rho_\Lambda}\right)^{\frac{1}{3}} \cdot senh^{\frac{2}{3}}\left[\sqrt{6\pi \cdot G \cdot \rho_\Lambda} \cdot t\right]$

En una primera etapa, posterior a la fase inflacionaria (fase que este modelo no describe) la expansión se produce desacelerada.

En una segunda etapa se acelera y finalmente la expansión se aproxima a una exponencial. (Figura 12.3).

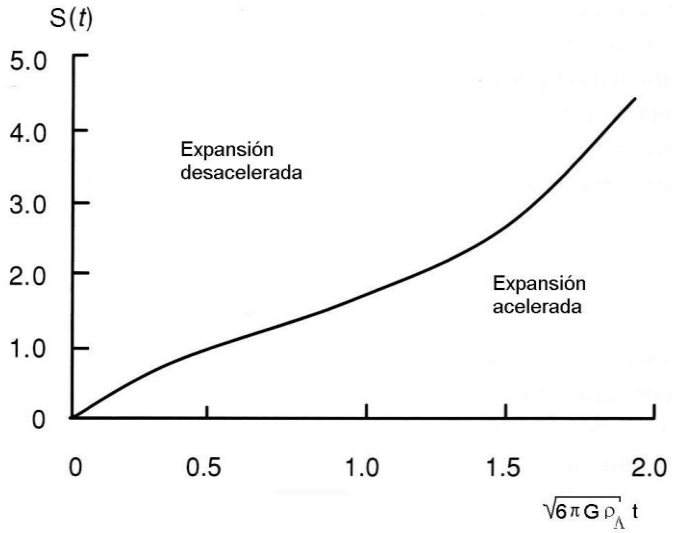


Figura 12.3 Evolución del factor de escala según el modelo estándar.

Si se pone $t = t_0$ en la fórmula para el factor de escala, éste vale 1 y se puede despejar una fórmula para t_0, a saber: $$t_0 = \frac{1}{\sqrt{6\pi \cdot G \cdot \rho_\Lambda}} \cdot arcsenh\left[\left(\frac{\rho_\Lambda}{\rho_m(t_0)}\right)^{\frac{1}{2}}\right]$$

Si el instante t_0 se toma en el presente y el origen de los tiempos en una singularidad inicial en el modelo matemático que corresponde al denominado Big Bang, entonces t_0 debe ser la edad actual del Universo, que el modelo predice a través de la fórmula que acabamos de deducir.

Valores para los parámetros del modelo cosmológico

Hay un cierto consenso tanto acerca de t_0, como acerca de los valores de los parámetros de densidad y acerca del parámetro de Hubble, a saber[215]:

$$t_0 = (13.7 \pm 0.2) \times 10^9 \cdot años \qquad \Omega_m(t_0) \approx 0.27 \pm 0.04 \qquad \Omega_r(t_0) \approx 0.00$$

$$\Omega_\Lambda(t_0) \approx 0.73 \pm 0.04 \qquad \frac{1}{r_0(t_0)} \cdot \frac{dr_0(t_0)}{dt} = H(t_o) = (70.4 \pm 1.5) \cdot \frac{km}{s \cdot Mpc}$$

Estos valores implican que el valor más verosímil del parámetro de curvatura ε es cero. De ser así, el Universo tendría una geometría plana, y la densidad total del fluido cósmico estaría próxima al valor crítico $\rho_{c,0} = \frac{3 \cdot H^2(t_0)}{8\pi \cdot G}$, valor que a su vez parece estar próximo a 1×10^{-26} kg/m^3.

Los modelos globales, basados en el principio cosmológico, intentan una descripción del comportamiento promedio del Universo, a gran escala, ignorando la riqueza de estructura y las heterogeneidades a escalas más pequeñas. En este sentido se parecen al modelo de geoide de la Tierra, que ignora los detalles del relieve cada lugar. Al igual que el geoide, suministran un escenario sobre el cual los detalles locales pueden insertarse y ser estudiados. Antes de considerar algunos de los aportes de la teoría de la relatividad al estudio de los detalles locales del cosmos, consideremos nuevamente el corrimiento al rojo del espectro de las galaxias lejanas.

12.4.3 Corrimiento al rojo cosmológico

La expansión del Universo produce un corrimiento al rojo del espectro observado de las galaxias que no se explica ni por el efecto Doppler de la relatividad restringida, estudiado en 8.4.1, ni por el efecto local de los campos gravitatorios, estudiado en 11.3.

Supongamos que un observador se encuentra en una posición radial fija $r_{observador} = 0$ respecto de un marco de referencia con la métrica de Friedmann-Lemaître-Robertson-Walker.

En un punto situado en $r_{fuente} = r_f$ existe una fuente que emite ondas luminosas que se dirigen hacia el observador y llegan a él.

---

[215] S. Serjeant "*Observational cosmology*", Cambridge University Press, U.K, 2010.
W. Freedman y E. Kolb "Cosmology" en G. Fraser (Ed.) "*The new physics in the twenty first century*", Cambridge University Press, U.K, 2009.

Supongamos que en un instante t_f la fuente emite la cresta de una onda, de longitud λ_f, y que un receptor junto al observador detecta es cresta en el instante t_r.
Pese a que el continuo tiempo-espacio-materia se idealiza como un medio homogéneo e isótropo, la métrica cambia con el paso del tiempo debido a que el factor de escala se modifica. Entonces la longitud de onda medida por el observador en el origen del marco de referencia es λ_r en vez de λ_f.
Para el movimiento radial de la luz se tiene:

$$ds^2 = c^2 \cdot dt^2 - r_0^2(t) \cdot A^2(\tilde{r}) \cdot d\tilde{r}^2 = 0 \qquad A(\tilde{r}) = \frac{1}{1 + \varepsilon \frac{\tilde{r}^2}{4}}$$

Entonces, separando variables, teniendo en cuenta que la luz viaja hacia el origen de coordenadas espaciales, e integrando resulta: $c \cdot \int_{t_f}^{t_r} \frac{dt}{r_0(t)} = -\int_{r_f}^{0} A(\tilde{r}) \cdot d\tilde{r}$

La siguiente cresta de onda es emitida por la fuente en el instante $t_f + \frac{\lambda_f}{c}$, una vez transcurrido un período de oscilación que en ese lugar es igual a $\frac{\lambda_f}{c}$. Esa nueva cresta es recibida en el instante $t_r + \frac{\lambda_r}{c}$ en el lugar donde se encuentra el observador. Entonces se tiene: $c \cdot \int_{t_f + \frac{\lambda_f}{c}}^{t_r + \frac{\lambda_r}{c}} \frac{dt}{r_0(t)} = -\int_{r_f}^{0} A(\tilde{r}) \cdot d\tilde{r}$ Restando miembro a miembro las ecuaciones para la primera y segunda cresta: $\int_{t_f + \frac{\lambda_f}{c}}^{t_r + \frac{\lambda_r}{c}} \frac{dt}{r_0(t)} - \int_{t_f}^{t_r} \frac{dt}{r_0(t)} = 0$ Esta ecuación se puede reescribir así: $\int_{t_r}^{t_r + \frac{\lambda_r}{c}} \frac{dt}{r_0(t)} = \int_{t_f}^{t_f + \frac{\lambda_f}{c}} \frac{dt}{r_0(t)}$ Ahora bien, como las integrales se extienden a intervalos de tiempo cuyas extensiones son iguales al período de la oscilación recibida y al período de la oscilación emitida, respectivamente, y como cabe suponer que el factor de escala permanece prácticamente constante durante esos intervalos de tiempo, se puede escribir:

$$c \cdot \int_{t_r}^{t_r + \frac{\lambda_r}{c}} \frac{dt}{r_0(t)} = \frac{\lambda_r}{r_0(t_r)} \qquad c \cdot \int_{t_f}^{t_f + \frac{\lambda_f}{c}} \frac{dt}{r_0(t)} = \frac{\lambda_f}{r_0(t_f)} \qquad \text{Entonces:} \quad \frac{\lambda_r}{r_0(t_r)} = \frac{\lambda_f}{r_0(t_f)}$$

Reordenado se obtiene para **el corrimiento** C **de origen cosmológico**:

$$1 + C = \frac{\lambda_r}{\lambda_f} = \frac{r_0(t_r)}{r_0(t_f)}$$

En un Universo en expansión $r_0(t_r)$ es mayor que $r_0(t_f)$, debido a lo cual el corrimiento debe ser al rojo. Además, cuanto más lejana se encuentra la fuente, mayor va a ser $r_0(t_r)$ respecto de $r_0(t_f)$, y más intenso va a resultar el corrimiento cosmológico.
Como lo resume Steven Weinberg: "El incremento en la longitud de onda desde la emisión hasta la absorción de la luz no depende de la velocidad de cambio del factor de escala cosmológico en los instantes de la emisión o de la absorción, sino del incremento del factor de escala en el período completo que va de la emisión a la absorción."[216]
De acuerdo con el modelo cosmológico estándar, se ve que el corrimiento al rojo C es el mismo para todas las longitudes de onda, y por tanto para todas las frecuencias, suponiendo los mismos instantes de emisión y de recepción.

12.5 El modelo cosmológico Λ-CDM y la teoría del Big Bang.

La suposición básica de los modelos cosmológicos revisados precedentemente es que el Universo está compuesto por una distribución regular, homogénea e isótropa de energía en forma de radiación y partículas de materia. En el mejor de los casos esta es una aproximación que se cumple en promedio a partir de una escala de 300×10^6 años-luz. Permite introducir un sistema de observadores en reposo respecto de la materia adyacente, en movimiento libre de fuerzas en un marco de referencia de Minkowski local. Para cada uno de estos observadores, las leyes físicas que se aplican en las correspondientes adyacencias son las que se verifican para el tiempo-espacio plano.
Estos observadores, que están en **co-movimiento** con la materia local, miden el tiempo local.
Se supone posible sincronizar los relojes de todos los observadores en co-movimiento mediante la siguiente regla: cuando observan que la densidad de la materia que los rodea toma determinado valor, entonces ajustan sus relojes para que se lea un determinado instante.
La simetría implícita en el postulado cosmológico asegura que los relojes permanezcan sincronizados. Se admite que con este procedimiento se puede definir un tiempo universal como el que aparece en la métrica de Friedmann-Lemaître-Robertson-Walker.

La distancia d entre dos observadores, cada uno en co-movimiento con la materia de su entorno, se define en la variedad de Riemann de curvatura variable con el transcurso del tiempo universal, y si la expansión es homogénea es isótropa entonces $d(t) = S(t)\cdot d(t_0)$ donde aparece el factor de escala cosmológico $S(t) = \dfrac{r_0(t)}{r_0(t_0)}$, el mismo para cualquier par de observadores.

Una analogía que ayuda a intuir la expansión del Universo consiste en imaginar un globo que se infla progresivamente. En la superficie del globo hay puntos marcados que representan las galaxias y cúmulos de galaxias. Podemos trazar una cuadrícula de coordenadas curvilíneas en la superficie del globo, de modo que los puntos que representan a cada una de las galaxias se encuentren en la intersección de dos líneas de coordenadas. A medida que se dilata la membrana elástica que forma la superficie del

[216] S. Weinberg "*Cosmology*", Oxford University Press, New York, 2008, p. 11.

globo, los puntos se separan, al igual que las líneas de coordenadas, pero sus coordenadas respecto de la red que se dilata no cambian.
Imaginemos ahora una población de hormigas en movimiento en la superficie del globo. Su movimiento representa el movimiento de las galaxias respecto de la velocidad local promedio de la materia.
Las hormigas poseen una velocidad máxima de desplazamiento respecto de la membrana elástica, de modo que cuando se cruzan dos hormigas en un mismo punto de la superficie del globo, la velocidad relativa de una hormiga respecto de otra no puede superar el doble de esta velocidad límite, estando ambos movimientos referidos a la membrana en expansión.
No obstante, si la expansión de la membrana es lo bastante rápida, dos hormigas situadas en puntos diferentes pueden estar separándose más rápido que el doble de la velocidad límite. Esto puede considerarse como una analogía con lo que ocurre en el Universo (considerado como un tiempo-espacio-materia en expansión), con las galaxias moviéndose con velocidades que no superan a la velocidad de la luz respecto del marco de referencia local.

El corrimiento al rojo del espectro de las galaxias junto con la Ley de Hubble, el comportamiento del brillo superficial[217] y la magnitud aparente de las galaxias distantes, junto con la radiación electromagnética considerada fósil, prácticamente isótropa, actualmente de cuerpo negro frío (con temperatura cercana a 3º K), se toman como evidencia de una expansión producida en grandes rasgos en forma acorde con el modelo estándar Λ-CDM de Universo.

Si se admite que el Universo puede haber tenido un comienzo, es concebible que ese comienzo haya sido una explosión gigantesca, conocida como Big Bang, a partir de lo que se conoce como una singularidad.
Durante los tres minutos siguientes a la gran explosión, se estima que nació la materia tal como la conocemos, en un proceso que se denomina núcleo-síntesis primordial.
El proceso de núcleo-síntesis fue propuesto por Gamow en 1942 y desarrollado por Alpher, Gamow y Herman a comienzos de la década de 1950[218].
En 1948, R. Alpher y R. Herman predijeron la existencia de una radiación de fondo de microondas asociada al Big Bang, que fue detectada recién en 1964.

Se supone que el Universo consistía al principio de una mezcla de corpúsculos de materia y antimateria cuya densidad y temperatura eran extraordinarias, en rápida expansión. Para impedir la aniquilación completa de la materia por la antimateria en una etapa posterior, más fría, parece razonable asumir, como lo hizo Gamow, que había un ligero exceso de materia respecto de la antimateria[219].

---

[217] El brillo superficial es flujo de energía proveniente de la galaxia, computado por unidad de ángulo sólido y por unidad de área donde se encuentra el observador. La magnitud aparente es una función del brillo superficial y tamaño angular aparente de la galaxia.

[218] R. Alpher "Una visión retrospectiva del modelo cosmológico del Big Bang" en J. Gonzalo, J. Sánchez y M. Alario (Eds.) "*Cosmología astrofísica*", Alianza, Madrid, 1995.

[219] Una interpretación alternativa puede hallarse en H. Alvén "Cosmology in the plasma Universe", IEEE Transactions on Plasma Science, volumen 18, 1990, pp. 5-10.

Los neutrinos se desacoplan del resto de la materia (el desacople se produce porque el intervalo de tiempo entre dos interacciones es mayor que la edad del Universo).
Los nucleones reaccionan parcialmente con los anti-nucleones por aniquilación recíproca y a medida que la expansión continúa la temperatura desciende. Los nucleones remanentes pueden comenzar a formar núcleos de helio y más pesados.
Los positrones reaccionan parcialmente con los electrones, mientras que los fotones resultantes de la aniquilación de materia con antimateria comienzan a interactuar mayoritariamente con los electrones remanentes. Las interacciones son ahora tan frecuentes que los fotones no se pueden propagar libremente y del intercambio de energía resulta un Universo cerca del equilibrio térmico: se forma un cuerpo aproximadamente negro.

Como vimos en 12.4.2, se suele aceptar que se puede realizar una descripción aproximada de las características a gran escala del Universo poniendo $\varepsilon = 0$ en la primera ecuación de Friedmann-Lemaître.
A partir de la simplificación introducida en 12.4.2, relacionado las densidades de materia-energía en cualquier instante inicial t_0 con sus valores en cualquier instante t a través del factor de escala, la dinámica viene entonces gobernada por la ecuación aproximada:

$$\frac{1}{r_0(t)} \cdot \frac{dr_0(t)}{dt} = \sqrt{\frac{8\pi \cdot G}{3} \cdot \left[\rho_m(t_0) \cdot \left(\frac{r_0(t_0)}{r_0(t)} \right)^3 + \rho_r(t_0) \cdot \left(\frac{r_0(t_0)}{r_0(t)} \right)^4 + \rho_\Lambda \right]}$$

Esta ecuación se puede integrar expresando el tiempo cósmico en función del radio de curvatura del tiempo -espacio-materia y de los parámetros de densidad introducidos en 12.4.2:

$$t(r_0) = \frac{1}{H_0} \cdot \int_0^{\frac{r_0}{r_0(t_0)}} \frac{dx}{x \cdot \sqrt{\Omega_\Lambda(t_0) + \Omega_m(t_0) \cdot x^{-3} + \Omega_r(t_0) \cdot x^{-4}}} \qquad H_0 = \sqrt{\frac{8\pi \cdot G}{3} \cdot \rho_c(t_0)}$$

Si se pone $r_0 = r_0(t_0)$ se obtiene una fórmula que se puede utilizar para calcular la edad actual t_0 del Universo. Si se desprecia el aporte del término asociado a $\Omega_r(t_0)$, se obtiene un modelo más simple (modelo Λ-CDM), con materia oscura fría y energía oscura dominante.
La introducción de la materia y la energía oscuras plantea varios problemas relacionados con la teoría cuántica de campos y la física de partículas elementales.
De momento no hay una explicación física satisfactoria para la energía oscura.
La materia oscura parece implicar la presencia de un tipo de partícula, reliquia de una etapa temprana de evolución del Universo, que todavía no ha sido detectada.
Por otro lado, no sabemos bien cómo serían las leyes físicas que se le podrían aplicar al estado de la materia en el comienzo mismo del Universo. Pero, no disponiendo de una alternativa mejor, podemos suponer que a partir de algún instante posterior al inicio del tiempo cósmico y a la denominada fase de inflación (que se considera más adelante), posiblemente desde algunos segundos a partir del Big Bang hasta la actualidad, se aplica este modelo matemático Λ-CDM.

Dado que al principio cabe esperar que el término correspondiente a la radiación haya sido enorme en comparación con los otros que aparecen en la cantidad sub-radical, la dinámica habría sido controlada por la radiación.

Como la densidad de radiación, proporcional a $\left(\frac{r_0(t_0)}{r_0(t)}\right)^4$ debería ser al comienzo enorme, si admitimos que la ley de Stefan según la cual la densidad total de radiación es proporcional a T^4, se aplica a esa radiación primigenia, entonces la temperatura $T(t)$ debería variar como $\frac{1}{r_0(t)}$ y ser extremadamente elevada.

Una vez que la expansión avanza lo suficiente, la radiación deja de dominar la dinámica, dando paso a un dominio de la materia primero y de la energía oscura después.

Cuando la radiación y la materia dominan la expansión, esta es desacelerada. Pero cuando comienza a dominar la energía oscura, la expansión se acelera.

Una vez transcurridos unos centenares de miles de años de tiempo cósmico, acompañando el descenso en la temperatura, los electrones se enlazan a los núcleos atómicos y los átomos comienzan a formarse. Las interacciones de los fotones se hacen mucho menos frecuentes, por lo cual se desacoplan y forman el fondo de radiación fósil isótropo y al parecer homogéneo, detectado por primera vez por en 1964 por Arno Penzias y Robert Wilson. Este hallazgo fue interpretado como una confirmación de la etapa de evolución del Universo que en el modelo de Alpher, Gamow y Herman se denominaba **bola de fuego**[220].

Posteriormente, nuevas observaciones y desarrollos teóricos condujeron a pensar que el origen de la radiación fósil debe hallarse en un gran destello originado en una emisión explosiva a partir de una generación precoz de estrellas hace unos 10×10^9 años, es decir unos 10×10^6 años contados a partir del inicio de la expansión del Universo [221].

Se supone que los elementos más pesados que el helio se sintetizaron en estas estrellas tempranas, antes de que se formaran las galaxias. Durante su fase de emisión explosiva, las estrellas precoces expulsaron gases que posteriormente se condensaron formando polvo cósmico.

El polvo cósmico, interactuando con los fotones, contribuyó a aproximarlos al equilibrio térmico (radiación de cuerpo negro). Luego la opacidad del universo habría desaparecido, los fotones se habrían podido propagar libremente y se habrían desacoplado lo suficiente de las partículas materiales como para que las características del espectro de radiación se mantuvieran de ahí en adelante, con una salvedad, que la temperatura de la radiación se fue modificando como consecuencia de la expansión del Universo.

La explicación del corrimiento al rojo de origen cosmológico, presentada en 12.4.3 y basada en el modelo cosmológico estándar, junto con consideraciones de termodinámica

---

[220] En el modelo cosmológico de Alpher, Gamow y Herman, el inicio de la expansión y la aparición de la bola de fuego prácticamente coinciden.

[221] Esta estimación del momento del destello se basa en asumir un corrimiento al rojo de 100, en la relación $1+C=\frac{\lambda_r}{\lambda_f}=\frac{r_0(t_r)}{r_0(t_f)}$ deducida en 12.4.3 y en la relación $r_0(t)\propto t^{\frac{2}{3}}$ correspondiente a un Universo dominado por la materia considerado en 12.4.2.

estadística permite describir la relación entre la temperatura de esta radiación de fondo y el factor de escala del Universo en cada instante de tiempo cósmico.
Admitiendo que la expansión del Universo produce un enfriamiento adiabático[222] de la radiación de cuerpo negro, cabría esperar que una consecuencia de la homogeneidad y la isotropía del cosmos sea que el **número relativo** de fotones de frecuencia f por unidad de volumen $\frac{n(f,T)}{n(f_0,T)}$ (donde $n(f,T)$ es la densidad de fotones correspondiente a una frecuencia f y f_0 es una frecuencia de referencia) se haya mantenido relativamente invariante, pero que la temperatura característica $T(t)$ de esta distribución haya venido disminuyendo conforme el radio de curvatura del Universo aumenta[223]:

$$T(t_2)\cdot r_0(t_2)=T(t_1)\cdot r_0(t_1)$$

Actualmente el fondo de radiación electromagnética fósil, prácticamente homogénea e isótropa presenta una temperatura extremadamente baja: 2.73º K.

En relación con la gran explosión y la subsecuente expansión del Universo, la homogeneidad e isotropía de la radiación de fondo medida por instrumentos a bordo del satélite COBE a partir de 1991, plantea un problema relacionado con la causalidad.
Para ver esto es preciso reconsiderar el horizonte de eventos de un suceso, es decir, los eventos más lejanos (y por ende más distantes hacia el pasado) que pueden influir sobre el suceso en cuestión.
En el momento en que las partículas de materia y la radiación se desacoplaron, este horizonte tenía una extensión (para los eventos de aquel entonces) muy reducida en comparación con la que tiene para los eventos presentes: apenas algunos grados de distancia medidos en la bóveda celeste. No obstante, la radiación fósil es como una fotografía transmitida hasta el presente de lo que ocurrió en aquel momento tan lejano en el tiempo, y muestra una sorprendente uniformidad que sugiere que las fuentes de la radiación se comportaron de manera igual aunque sus horizontes de eventos en ese momento no les permitiera interactuar intercambiando información.

Una posible explicación es el modelo de inflación de Guth[224] . Guth conjeturó que en los primeros instantes de vida del Universo la densidad estaba dominada por un campo global, **el inflatón**, que se asocia a una dilatación súbita tiempo-espacio-energía (a la cual que denominó inflación), en un factor de 10^{40}. La duración del proceso inflacionario es incierta, pero debe haber sido muy breve. (Figura 12.4)

---

[222] R. Tolman "*Relativity, thermodynamics and cosmology*", Clarendon Press, Oxford, 1934.

[223] Una demostración cuidadosa puede encontrarse en E. Peebles "*The large structure of the Universe*", Princeton University Press, Princeton, 1980.

[224] E. Harrison "*Cosmology: the science of the Universe*", Cambridge University Press, Londres, 2008.
J. Rich "*Fundamentals of Cosmology*", Springer Verlag, Heildelberg, 2001.

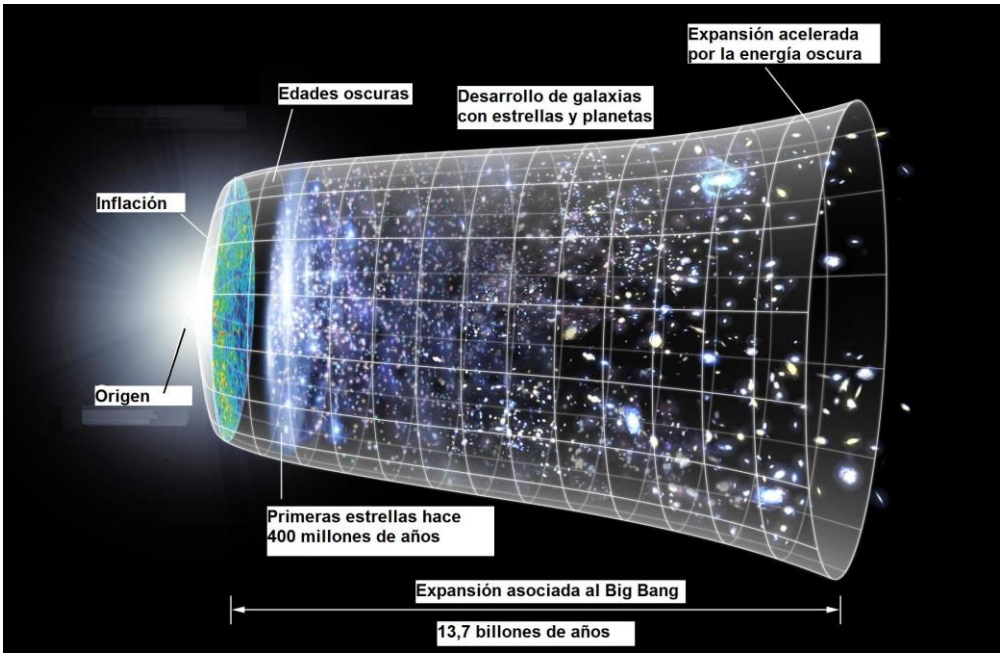


Figura 12.4 Representación de las etapas de evolución del Universo de acuerdo con el modelo estándar, incluyendo la fase inflacionaria inicial. El espacio cósmico aparece, para cada instante de tiempo cósmico, como una circunferencia ortogonal al eje del tiempo. Al comienzo la expansión presenta una enorme aceleración asociada al proceso de inflación. Luego se desacelera hasta que, más cerca del presente, la expansión del Universo comienza a acelerarse nuevamente.

En principio, la inflación podría explicar parcialmente algunas características del Universo que el modelo original del Big Bang no puede explicar y debe tener en cuenta bajo la forma de postulados: el elevado grado de aproximación a una geometría plana que parece presentar el espacio físico (las hiper-superficies espacialoides del flujo de Hubble), por qué el Universo parece ser homogéneo e isótropo a gran escala pero heterogéneo a escalas menores.

Se supone que cuando la temperatura descendió lo suficiente la influencia del inflatón se desvaneció y comenzó la expansión propiamente dicha, dominada al principio por la radiación. Una vez que la dinámica del Universo pasa de ser gobernada por la radiación a ser gobernada por la materia, se habrían formado las galaxias y continuado la expansión.

Se supone que al salir de la etapa de inflación la materia presentaba fluctuaciones microscópicas de densidad: los gérmenes de las futuras galaxias. Cabría esperar que esas fluctuaciones se traducirían en fluctuaciones, pequeñas pero observables, en el fondo de radiación fósil.

A partir de 1992 se pudo detectar una anisotropía muy pequeña pero significativa[225] en la intensidad de la radiación proveniente de diferentes direcciones que incidía sobre el satélite COBE.

Se interpreta como evidencia experimental de la formación de los gérmenes de las actuales galaxias.

[225] Empleando un instrumento conocido como radiómetro diferencial de microondas, se midieron fluctuaciones en la radiación con valores (relativos al valor promedio) del orden de 10^{-5}.

12.6 Modelo de Universo en mosaico.

Los modelos basados en la métrica estándar no permiten tener en cuenta la estructura discreta que se observa en pequeña escala. Esta estructura se caracteriza por conjuntos de astros con interacción gravitatoria mutua significativamente más fuerte que la interacción que los astros del conjunto presentan con conjuntos más lejanos. En la medida en que esto ocurre a cada uno de estos conjuntos se le puede denominar sistema auto-gravitante.

Un modelo un poco más realista, muy simple pero que tiene en cuenta la estructura heterogénea formada por los sistemas auto-gravitantes, es el **modelo de mosaico**.

Consideremos un caso en el cual la densidad de masa es uniforme, excepto en una pequeña región esférica cuyo radio es proporcional al factor de escala $r_0(t)$ y supongamos que la totalidad de la masa necesaria para que esta región posea la misma densidad que el resto del Universo se encuentra concentrada en el centro de la esfera. El resto forma una cavidad hueca. Ahora bien, el campo gravitatorio fuera de la cavidad es el mismo que se hallaría si la masa puntual se encontrara uniformemente distribuida. Como consecuencia la presencia de la cavidad no modifica la dinámica global del modelo de Universo: se expande igual que si la cavidad no existiera. Pero además, el resto del Universo no incide sobre el campo gravitatorio en el interior de la cavidad. Este último viene dado por la métrica de Schwarzschild, como si la masa puntual que ocupa el centro de la cavidad se encontrara en el espacio vacío. No obstante, en la frontera de la cavidad las estructuras de tiempo-espacio asociadas a las dos métricas deben coincidir. Como el resto del Universo no rota localmente, la masa puntual tampoco debe rotar.

Es posible sustituir la masa puntual por un sistema auto-gravitante lo bastante pequeño. No es necesario suponer que la distribución de la densidad en el exterior de la esfera sea exactamente uniforme, lo cual nos permite contemplar la presencia de otros agrupamientos de astros análogos al que ocupa la parte central de la esfera. La única condición que debe cumplirse es la coincidencia de las estructuras de tiempo-espacio en las fronteras.

Como los efectos gravitatorios de los agrupamientos de astros no se extienden mucho más allá de distancias del mismo orden numérico que la dimensión de esos agrupamientos, la dinámica del Universo en sí mismo, su estructura de tiempo y espacio en gran escala, es el resultado de su contenido de materia (partículas y radiación), pero no influye directamente en la dinámica en pequeña escala de los sistemas auto-gravitantes, incluidas las galaxias. Suministra un conjunto de marcos de referencia locales que no se encuentran acelerados respecto del marco de referencia cósmico.

Este modelo plantea un Universo en mosaico, en el cual los agrupamientos locales distorsionan débilmente la estructura del tiempo-espacio, de modo que se puede aplicar localmente el tipo de aproximación utilizada en el estudio de las lentes gravitacionales o del corrimiento al rojo de la luz en los campos gravitatorios considerados en 11.6 y 11.3, respectivamente.

13. La metagalaxia

"Al aumentar la distancia, nuestro conocimiento se desvanece, y se desvanece rápidamente. Eventualmente, alcanzamos las dimensiones fronterizas-los límites de los telescopios. Allí medimos sombras, y buscamos entre errores fantasmales de medida, señales que son apenas más sustanciales que estos errores"
Edwin Hubble, "The realm of the nebulae", Yale University Press, New Haven, 1936.

"Nosotros no somos más que polvo de estrellas, cada elemento que nos compone contiene carbono, oxígeno y otros elementos que han sido todos, un día, sintetizados en una estrella, que muriendo en mil fuegos, esparció la simiente necesaria para nuestra existencia"
Hubert Reeves, en "Poussieres d'étoiles", Seuil, Paris, 1994.

"Chandrasekhar muestra que una estrella cuya masa supera un cierto límite permanece como un gas perfecto y no puede enfriarse hasta el equilibrio. La estrella tiene que continuar radiando y radiando y contrayéndose y contrayéndose hasta que, supongo, alcanza un radio de unos pocos kilómetros cuando la gravedad se hace lo bastante intensa como para retener la radiación y la estrella puede al fin hallar la paz.
Me siento impelido a concluir que esto es casi una reducción al absurdo de la fórmula de degeneración relativista.
Varios accidentes pueden intervenir para salvar la estrella, pero yo quiero más protección que esa.
Pienso que debería haber una ley de la naturaleza que impida a la estrella comportarse de esta manera absurda".
Arthur Eddington comentando un trabajo de su alumno Subrahmanyan Chandresakhar, presentado en enero de 1935 en una reunión anual de la Royal Astronomical Society.

"Cuando discutimos las variadas posibilidades que aparecen como resultado de las interacciones con agujeros negros o entre agujeros negros, estamos hoy día considerando seriamente situaciones que hace no mucho tiempo fueron barridas hacia un costado como reducciones al absurdo.
Por mi parte, mientras considero los fenómenos asociados con los horizontes de eventos y la imposibilidad de comunicación a través de ellos, a menudo recuerdo una parábola sobre la naturaleza que aprendí hace cincuenta años en la India.
La parábola se refiere a larvas de libélula depositadas en el fondo de un estanque.
Una fuente constante de misterio para esas larvas es lo que ocurre, una vez alcanzado el estadio de crisálida, cuando pasan a través de la superficie del estanque para no regresar jamás.
Cada larva, cuando se acerca el momento en el que se transformará en crisálida, promete regresar y contar a las que permanecieron en el estanque lo que realmente ocurre más allá, negando o confirmando un rumor, atribuido a una rana, según el cual cuando una larva emerge de la superficie del estanque, al otro lado de su mundo, deviene una criatura maravillosa, con un cuerpo largo y delgado y alas iridiscentes.
Pero cuando emerge del estanque como libélula completamente formada, resulta incapaz de retornar, de penetrar la superficie del agua, pese a todos sus esfuerzos y por más tiempo que dedique.
Así, los libros de historia de las larvas no registran regreso alguno para contar lo que les ocurre al atravesar el domo de su mundo. Y la parábola finaliza con una súplica: ¿Ninguno de ustedes, por piedad, a aquellos dejados atrás, les develará el secreto?"

Subrahmanyan Chandresakhar, en una conferencia sobre el rol creciente de la teoría generalizada de la relatividad en las investigaciones astronómicas y astrofísicas, en la Universidad de Oxford, en mayo de 1972.

La cita de Hubble al comienzo de este capítulo llama la atención sobre un aspecto importante relacionado con las incertidumbres en las mediciones astronómicas y astrofísicas. Pese a los notables progresos en la instrumentación y pese a las posibilidades que brinda la astronáutica, a través de las nuevas plataformas de observación (las sondas espaciales) disponibles para instalar y operar los instrumentos significativamente mejorados, la dificultad de fondo que señala Hubble permanece cuando el interés de los estudiosos no se limita ya a una región relativamente acotada, como el Sistema Solar, sino abarca la totalidad del cosmos. Cuando se trata de elegir entre modelos cosmológicos sobre la base de evidencia empírica, aparece un escollo fundamental: la inmensa duración del intervalo de tiempo que al parecer se requiere para acumular datos en calidad y cantidad adecuada para someter a teorías, a menudo antagónicas, a la necesaria confrontación con la experiencia. Entonces, si no se abandona el intento, se corre el riesgo de recurrir a razones de índole religiosa, filosófica, e inclusive estética. La cita de Eddington, opinando sobre el trabajo de Chandrasekhar, es bastante elocuente al respecto.

Todo lo que podemos observar desde la Tierra, es una esfera (formada por líneas geodésicas que parten desde el observador) de un radio aproximado de 46×10^{9} años-luz. Se estima que en ese **universo observable**, denominado **Metagalaxia**, hay 10^{11} galaxias.
El diámetro de nuestra galaxia es de solo 1×10^{5} años-luz. Una distancia común entre galaxias próximas es 3×10^{6} años-luz.
Habría un número de estrellas del orden de 10^{23} en la Metagalaxia.
A escalas cuyos órdenes de magnitud son inferiores a 300×10^{6} años-luz se tiene evidencia de una estructura jerárquica. La mayor parte de los átomos y de la radiación forman estrellas, aproximadamente esféricas y en equilibrio fluido-dinámico. La mayor parte de las estrellas forman galaxias y las galaxias forman cúmulos de galaxias. Los cúmulos se incorporan en súper-cúmulos y éstos en grandes estructuras, hasta la escala en la que se asume homogeneidad e isotropía y se aplica el principio cosmológico.
La densidad promedio se estima en 9.9×10^{-30} gramos por centímetro cúbico, con un 74% de energía oscura, un 22% de materia oscura fría y un 4% de materia ordinaria. En promedio habría el equivalente en masa a un átomo de hidrógeno cada 4 metros cúbicos. No parece haber carga eléctrica neta a escala cosmológica, lo que hace que la gravitación domine la dinámica a gran escala.
Las estrellas nacen, evolucionan y mueren. La energía que irradian procede de reacciones de fusión de los núcleos atómicos del material del interior estelar, a temperaturas de millones o decenas de millones de grados.
Aproximadamente nueve décimas partes de la masa de ese material es hidrógeno y helio. Los elementos más pesados, que se encuentran en distintas partes de la Metagalaxia, y en particular en nuestro rincón, en la Tierra y la Luna, se formaron en las entrañas de las estrellas.
Se plantea entonces el problema de la génesis, evolución y transformación de la materia para formar las estrellas y las galaxias.

El carácter irregular, más bien de tipo fractal, (en cúmulos, super-cúmulos, super-super-cúmulos,…de galaxias) que presenta la distribución de la materia a escalas relativamente pequeñas puede tenerse en cuenta mediante un **principio cosmológico estadístico**[226]: la distribución de materia sigue las mismas leyes estadísticas con independencia del marco de referencia empleado para la observación. Si no se establece un corte a una escala lo suficientemente grande, se obtienen sistemas cuyas masas y diámetros superan cualquier límite pre-establecido y cuyas densidades se hacen cada vez menores, tendiendo a cero: este es el modelo de Universo estacionario que C. Charlier propuso en 1908.

El empleo de fractales para describir la estructura del Universo originó una clase de teorías cosmológicas conocidas colectivamente como "**cosmología fractal**".
A partir de 1989, L. Nottale, acompañado por un pequeño grupo de investigadores, asumiendo un modelo continuo pero no diferenciable del tiempo-espacio, extendió el principio de relatividad para tener en cuenta la invariancia de las leyes frente a cambios en la escala en la que se describen las coordenadas de un evento[227].
Por el momento y en su forma actual, la denominada Teoría de la Relatividad de Escala (TRE) ocupa una posición marginal dentro del ámbito de la física. A diferencia de lo que ocurre con teorías de vanguardia como las teorías de super-cuerdas y las especulaciones sobre la materia oscura, los resultados de la TRE son ignorados por la mayor parte de los físicos y en consecuencia no se puede considerar ciencia aceptada.

13.1 Modelos de Metagalaxia

Algunas personas consideran inadecuado el uso de modelos cosmológicos que describen partes de un Universo que no es ni será observable.
Por este motivo prefieren modelar la Metagalaxia, considerado como un sistema de galaxias **auto-gravitante máximo**. En general admiten que hay una escala máxima para la distribución cuasi-fractal de la materia, y que a escalas mayores se puede recurrir a modelar la Metagalaxia asumiendo homogeneidad e isotropía en promedio: se aplica entonces el mismo principio cosmológico utilizado en los modelos de Universo.
Lo más simple que puede hacerse en este caso es partir de un modelo como, por ejemplo, el hiperbólico o el plano, agregando una frontera en la coordenada radial.
En ese caso el volumen $V_{Meta}(t)$ y la masa M_{Meta} están acotados, mientras que en el modelo original de Universo hiperbólico o plano el volumen es ciertamente infinito.
Si suponemos un modelo de Metagalaxia hiperbólico caracterizado porque la coordenada radial está acotada $r \leq r_F$, el volumen se puede hallar mediante la fórmula presentada en 12.4.1 al final de la parte sobre el modelo hiperbólico. Si se supone un modelo plano, el volumen es el de una esfera de radio r_F. En ambos casos r_F se expande en forma proporcional al factor de escala cósmico.
Si $\rho_{Meta}(t)$ es la densidad promedio, la masa M_{Meta} de la Metagalaxia se puede estimar mediante la fórmula $M_{Meta} = \rho_{Meta}(t) \cdot V_{Meta}(t)$

[226] B. Mandelbrot "*La geometría fractálica de la naturaleza*", Taurus, Madrid, 1998.

[227] Un resumen sobre los controvertidos modelos fractales del tiempo-espacio y los desarrollos basados en la teoría de la relatividad de escala se puede hallar en el artículo de L. Nottale "Scale relativity and fractal space-time: theory and applications", Foundations of Science, vol. 15, Nº 2(3), 2010, pp. 101-152.

Si bien el modelo hiperbólico no acotado y el modelo hiperbólico acotado no difieren en sus predicciones en la medida en que nos mantengamos lejos del borde del modelo acotado y no retrocedamos mucho hacia el pasado, la diferencia puede ser significativa cuando estas condiciones no se verifican[228].
Suponiendo que el volumen $V_{Meta}(t)$ es proporcional al cubo de un radio característico $R_{Meta}(t)$, y que en ningún momento de la evolución de la Metagalaxia ese radio puede haber sido inferior al radio de Schwarzschild correspondiente a la masa de la Metagalaxia $r_{S,Meta} = \frac{2 \cdot G \cdot M_{Meta}}{c^2}$, de esta última condición y de la relación $M_{Meta} = \rho_{Meta}(t) \cdot V_{Meta}(t) \approx \rho_{Meta}(t) \cdot \frac{4\pi}{3} \cdot R^3_{Meta}(t)$ se desprende un límite superior para la densidad promedio: $\rho_{Meta}(t) \leq \frac{3 \cdot c^6}{32\pi \cdot G^3 \cdot M_{Meta}}$
Por ejemplo, si M_{Meta} fuera del orden de 10^{54} g, entonces el límite de densidad sería del orden de los 10^{-25} g/cm^3. En el estadio llamado de la bola de fuego, la densidad promedio del Universo primitivo se supone tomaba valores del orden de 10^4 g/cm^3, claramente incompatible con las hipótesis de una masa constante y de dimensiones acotadas inferiormente por un radio de Schwarzschild.
Este tipo de consideración, junto con argumentos de simetría en relación con la proporción de materia y antimateria, condujeron a O. Klein y a su alumno H. Alfvén a desarrollar un modelo de nube primordial de materia (protones) y antimateria (antiprotones) en partes iguales, el ambiplasma, extremadamente tenue, que se habría contraído por efecto del campo gravitatorio. Las leyes de la magneto-hidrodinámica resultan fundamentales para comprender lo que acontece según este modelo[229].

El modelo cosmológico plano sirvió de base para una explicación de la formación de las galaxias, conocida como hipótesis de acumulación gravitatoria[230], alternativa a la hipótesis de la bola de fuego. La hipótesis de la acumulación gravitatoria originó una teoría que da cuenta de la estructura jerárquica de las agrupaciones estelares. Permite explicar la naturaleza fractal de esta estructura hasta alcanzar una dimensión de corte de aproximadamente 300×10^6 años-luz.

Resulta interesante que al haber introducido una acotación en la coordenada radial del modelo hiperbólico o en la del modelo plano, el marco de espacio que éste suministra queda físicamente limitado, pero esta limitación no se corresponde con un límite propio

---

[228] Varias consideraciones sobre el modelo hiperbólico de Metagalaxia, basadas en la evidencia disponible en aquel entonces, pueden hallarse en O. Klein "Arguments concerning relativity and cosmology", Science, vol. 171, 1971, pp. 339-345.

[229] Una buena descripción de estas ideas de Klein y Alfvén se puede hallar en el artículo de H. Alfvén "Antimatter and cosmology", Scientific American, abril de 1967.

[230] Una exposición muy clara del mecanismo de acumulación gravitatoria y de la evidencia empírica que motivó la introducción de la hipótesis de una generación de estrellas pre-galácticas y el gran destello que les puso fin, realizada por la persona que propuso ese mecanismo ya en 1951, se puede encontrar en D. Layzer "*Construcción del Universo*", Labor, Barcelona, 1989.

del modelo matemático, puesto que ni el modelo hiperbólico ni el modelo plano lo tienen. Esto constituye una diferencia muy significativa respecto del modelo esférico.

13.2 Estrellas

Una estrella es un cuerpo formado por un gigantesco conglomerado gaseoso de partículas y radiación, confinado por su propio campo gravitatorio, que equilibra el efecto disgregante de la presión de las partículas del gas y de la presión de la radiación.
En forma algo sobre-simplificada puede decirse que una estrella nace cuando una masa lo bastante grande de gas compuesto por una mezcla en las que el hidrógeno y el helio son mayoritarios[231], se contrae bajo la acción de la gravedad. A medida que la masa se contrae, un núcleo central se calienta y la presión aumenta. Cuando la presión y la temperatura han aumentado lo suficiente, se desencadena un proceso de fusión de núcleos de hidrógeno[232].
Las estrellas pueden mantener un estado prácticamente estacionario durante miles de millones de años, liberando energía termonuclear de fusión en su región central (que se encuentra por lo general a temperaturas del orden de los millones o decenas de millones de grados) y radiando la potencia producida hacia el espacio circundante.
En el caso del Sol, compuesto de 71% de hidrógeno, 27% de helio, y de pequeños porcentajes de otros elementos (carbono, nitrógeno, oxígeno, neón, magnesio, silicio, azufre y hierro), cuatro núcleos de hidrógeno reaccionan en etapas para dar lugar a un núcleo de helio. Como la suma de las masas de reposo de los protones m_p es superior a la masa de reposo del núcleo de helio m_{He}, durante ese proceso se libera una energía $\left(4 \cdot m_p - m_{He}\right) \cdot c^2$, igual a 27 MeV. Esta es la fuente de la energía que el Sol radia a razón de 4×10^{23} kW. Esta potencia radiada se denomina **luminosidad**.
La masa solar estimada es de 2×10^{30} kg, el radio solar es de 6.9×10^{5} km y su edad actual se estima en 5×10^{9} años.
En su fase actual, se estima que la luminosidad del Sol debe permanecer prácticamente constante durante miles de millones de años.
Cuando se consume el hidrógeno, la estrella se contrae, aumentando la temperatura en su centro hasta alcanzar un valor crítico que hace posible un nuevo tipo de reacción de fusión liberadora de energía: tres núcleos de helio originan un núcleo de carbono y se establece un nuevo estado estacionario. Cuando se consume el helio, la estrella comienza a contraerse de nuevo, aumentando aún más su temperatura interna, hasta alcanzar un nuevo valor crítico que hace posible otra reacción de fusión nuclear exotérmica.
Comparando este proceso de fusión nuclear, lenta y estable, con la brusca detonación de las bombas de fusión que conocemos bien aquí en la Tierra, cabe preguntarse a qué se debe la estabilidad de las estrellas.
El teorema del virial, visto en 11.6 en relación con la materia oscura cósmica, puede utilizarse para analizar lo que ocurre en una estrella como el Sol, en condiciones en que se pueden despreciar los efectos de la presión de los corpúsculos sobre la energía total de la estrella. Cuando se puede despreciar los efectos de la presión, la energía total de

---

[231] El 98 % de la masa de la Vía Láctea está compuesta de hidrógeno y de helio. Los elementos más pesados son producidos por reacciones de fusión nuclear en el interior de las estrellas.

[232] Una descripción verbal detallada del nacimiento, evolución y desaparición de las estrellas puede hallarse en el libro de J. Comellas "*El mundo de las estrellas*", Equipo Sirius, Madrid, 1999.

una estrella se puede expresar como suma de la energía cinética E_c de las partículas y la energía potencial gravitacional del sistema Φ: $E_\oplus = E_c + \Phi$

El teorema del virial asegura que $2 \cdot E_c + \Phi = 0$. Supongamos que esta igualdad se mantiene, al menos con buena aproximación, cuando la estrella se aparta ligeramente de su estado estacionario. Entonces, del teorema del virial resulta $E_\oplus = E_c + \Phi = -E_c$.

Pero, por otro lado, $E_c \approx \frac{3}{2} \cdot N \cdot k_B \cdot \langle T \rangle$ donde N es el número de partículas en el gas estelar, k_B es la constante de Boltzmann y $\langle T \rangle$ es la temperatura promedio de la estrella.

Si L es la luminosidad y F es la potencia liberada por fusión termonuclear, entonces la velocidad de variación de la energía total de la estrella viene dada por: $\frac{d}{dt} E_\oplus = F - L$

En estado estacionario $F = L$ y como consecuencia $E_\oplus$ permanece constante.

Supongamos que, debido a una perturbación, $F > L$ Entonces $\frac{d}{dt} E_\oplus > 0$ por lo cual $E_\oplus$ debe aumentar. Pero como $E_\oplus = -E_c$, resulta que $\frac{d}{dt} E_c < 0$ con lo cual E_c debe disminuir. Si disminuye E_c, disminuye $\langle T \rangle$, y si disminuye $\langle T \rangle$ esto trae como consecuencia una disminución de F. Esta disminución se mantiene mientras se verifique $F > L$, así que esta secuencia de procesos tiende a conducir a la estrella a un estado estacionario.

Si $F < L$ la energía $E_\oplus$ disminuye, E_c aumenta al igual que $\langle T \rangle$, lo que trae como consecuencia un aumento en F que continúa mientras se verifique $F < L$. La secuencia de procesos tiende nuevamente a conducir a la estrella a un estado estacionario.

Así pues, un sistema termonuclear auto-gravitante posee una regulación automática de la velocidad con que se desarrollan las reacciones de fusión en su interior, basada en un **calor específico negativo**: si su energía **aumenta** (disminuye) su temperatura **disminuye** (aumenta).

13.2.1 Enanas blancas

Cuando se agotan las posibilidades de nuevas reacciones de fusión exotérmicas, la estrella continúa disipando energía, lo cual disminuye su temperatura. Al disminuir la temperatura la presión del gas no puede impedir el colapso gravitatorio de la estrella.

Si esta posee una masa inferior a un valor límite de aproximadamente 1.4 veces la masa del Sol, se contrae hasta formar una especie de átomo gigantesco.

Los núcleos han perdido sus electrones (al menos los de los orbitales externos). Los electrones forman un gas al que se aplica el principio de exclusión de Pauli: dos electrones no pueden tener los mismos números cuánticos. Como el número de electrones es enorme, se llenan todos los estados cuánticos desde el de mínima energía hasta estados cuyas energías son considerables.

Como consecuencia de este efecto mecánico-cuántico, la agitación y la presión son muy grandes.

Aún a bajas temperaturas, cercanas a los 0º K, la presión del gas de electrones es capaz de detener el colapso gravitatorio, pero lo puede hacer en condiciones en las que las

densidades estelares son muy grandes, del orden de 10^8 g/cm^3 y el radio estelar muy pequeño, del orden de los 1000 km[233].
A estos astros se los conoce como **enanas blancas**. Son poco luminosas porque tienen poca energía para radiar.
Los núcleos atómicos no contribuyen en forma significativa a la presión en el interior de las enanas blancas[234].
Cuando se comenzó a investigar el proceso de enfriamiento de las estrellas con las herramientas que suministraba la mecánica cuántica, se construyeron modelos **no relativistas** de la dinámica electrónica. Se basaban en en una ecuación de estado propuesta por R. Fowler en 1926 para describir la relación entre presión, densidad y temperatura de un gas electrónico. Este enfoque del problema del enfriamiento estelar condujo a predecir una relación entre la masa M del astro y su radio R en el estado de equilibrio del tipo $M = M_* \times \left(\frac{R_*}{R}\right)^3$ donde M_* y R_* son constantes. De acuerdo con esta relación, el radio de la estrella disminuye con el aumento de la masa, pero en principio son posibles estados de equilibrio para todo valor de la masa estelar. Aún las estrellas más masivas podrían enfriarse en forma estable hasta un estado de equilibrio final. Pero los cálculos de las cantidades de movimiento de los electrones en el centro de las estrellas ya enfriadas sugerían que debían tomar valores comparables con $m_e \cdot c$, donde m_e es la masa del electrón. En estas condiciones la dinámica es relativista, no clásica, de modo que se hizo evidente la necesidad de modificar los modelos de enfriamiento estelar.
El trabajo pionero de Chandrasekhar, en 1935, introdujo la dinámica de la relatividad restringida en los modelos de evolución estelar. Los resultados del nuevo modelo sugerían que la evolución de las estrellas de masa pequeña debía ser esencialmente diferente de la evolución de las estrellas muy masivas. Mientras que para una estrella de masa pequeña el estadio de enana blanca parece ser el paso previo a la extinción completa, una estrella masiva no puede pasar sin más al estadio de enana blanca. Es necesario contemplar otras posibilidades.
La interpretación de datos experimentales, obtenidos posteriormente, confirma las predicciones de los modelos relativistas, incluyendo un estadio alternativo al de enana blanca: la estrella de neutrones y su manifestación electromagnética, el pulsar.
Sin entrar en los detalles del trabajo de Chandrasekhar sobre las enanas blancas y de trabajos posteriores de otros autores sobre las enanas blancas y las estrellas de neutrones, como los de J. Oppenheimer, G. Volkoff y L. Landau entre otros, se pueden construir modelos relativistas muy simple de enanas blancas y de estrellas de neutrones. Tanto para las enanas blancas como para las estrellas de neutrones, el modelo simplificado permite estimar sus radios y predecir la existencia de masas umbral, por encima de la cual el estado de la materia característico de estos estadios de la evolución estelar no es estable.

---

[233] La densidad promedio del Sol es de 1.41 g/cm^3 y su radio del orden de los 10^5 km. La densidad promedio de la Tierra o de Venus es de 5.5 g/cm^3.

[234] Las propiedades de la materia a densidades elevadas pueden estudiarse en el capítulo 11 del libro de L. Landau y E. Lifchitz "*Física estadística*", Reverté, Barcelona, 1982.

Para construir un modelo relativista de enana blanca, se comienza por una expresión de la energía $E_{\oplus}$ del sistema como suma de un término que tiene en cuenta la combinación de la energía cinética con la energía de reposo para los electrones y un término que tiene en cuenta la energía potencial debida a la interacción gravitatoria entre los núcleos que forman el sistema. Se desprecia el aporte de la agitación de los núcleos y la energía potencial debida a la interacción electroestática entre las cargas, asumiendo que los términos de repulsión equilibran a los términos de atracción en el interior del plasma estelar, debido a que carga neta es nula y los núcleos atómicos y los electrones se encuentran íntimamente entremezclados.

Consideremos un sistema formado por N átomos de número atómico Z y número de masa A, completamente ionizados.

Teniendo en cuenta que, como se mostró al final de 7.1, la energía E de una partícula con cantidad de de movimiento p y masa en reposo m_0 se puede expresar mediante la fórmula $E = \sqrt{m_0^2 \cdot c^4 + c^2 \cdot p^2}$, el término debido a la energía de los electrones se puede expresar mediante $N \cdot Z \cdot \sqrt{m_e^2 \cdot c^4 + c^2 \cdot p^2}$. En esta expresión m_e es la masa en reposo del electrón y p es una cantidad de movimiento promedio.

El término de interacción gravitatoria entre los núcleos se puede estimar mediante la expresión $-\frac{1}{2} \cdot N^2 \cdot G \cdot \frac{(A \cdot M_n)^2}{R}$. En esta expresión M_n es la masa en reposo del nucleón, sin distinguir a estos efectos protones de neutrones, y R es el radio del sistema, sistema que por simetría se asume adopta una forma esférica.

Entonces la energía total es: $E_{\oplus} = N \cdot Z \cdot \sqrt{m_e^2 \cdot c^4 + c^2 \cdot p^2} - \frac{1}{2} \cdot N^2 \cdot G \cdot \frac{(A \cdot M_n)^2}{R}$

En esta fórmula $A \cdot M_n$ representa la masa de los núcleos atómicos.

Si el sistema posee un radio R su volumen es proporcional a R^3. Como a los electrones se les aplica el principio de exclusión de Pauli, podemos suponer que a cada uno le corresponde un volumen proporcional a $\frac{R^3}{N \cdot Z}$. Entonces, combinando esta estimación con el principio de incertidumbre de Heisenberg para el par de variables conjugadas posición y cantidad de movimiento se obtiene[235]: $\frac{R}{(N \cdot Z)^{\frac{1}{3}}} \cdot p \approx \hbar$

Introduciendo $\frac{1}{R} \approx \frac{c \cdot p}{\hbar \cdot c \cdot (N \cdot Z)^{\frac{1}{3}}}$ en la fórmula para la energía, dividiendo por el número total de electrones y reordenado se obtiene la siguiente fórmula para la energía total electrón:

$$e_e(p) = \frac{E_{\oplus}}{N \cdot Z} = \sqrt{m_e^2 \cdot c^4 + c^2 \cdot p^2} - \left(\frac{N}{N_{cr}}\right)^{\frac{2}{3}} \cdot c \cdot p$$

[235] Como es usual, $\hbar$ representa el cociente entre la constante de Planck h y 2π.

Aquí por definición se ha puesto $N_{cr} = N_0 \cdot \left(\frac{Z^2}{A^3}\right)$ con $N_0 = \left(\frac{2}{\alpha_G}\right)^{\frac{3}{2}}$ y $\alpha_G = \frac{G \cdot M_n^2}{c \cdot \hbar}$.

El número sin dimensiones $\alpha_G \approx 6.3 \times 10^{-39}$ se conoce como **constante de interacción gravitatoria**. El cálculo de $N_0 = \left(\frac{2}{\alpha_G}\right)^{\frac{3}{2}}$ da 6×10^{57}.

Asumimos que el estado que adopta el sistema es el que minimiza la energía $e_e(p)$ como función de p (y por ende, minimiza la energía total $E_\oplus$, puesto que el número total de núcleos N viene dado y no varía).

Puesto que $\frac{d^2}{dp^2} e_e(p) > 0$ para todo p, el mínimo se presenta cuando $\frac{d}{dp} e_e(p) = 0$.

Resolviendo esta última ecuación se obtiene que $\frac{c \cdot p}{\sqrt{m_e^2 \cdot c^4 + c^2 \cdot p^2}} = \left(\frac{N}{N_{cr}}\right)^{\frac{2}{3}}$

Como el miembro de la izquierda de esta igualdad es inferior a 1, se desprende que existe una solución siempre que el número total de átomos de la estrella no supere el valor crítico $N_{cr} = N_0 \cdot \left(\frac{Z^2}{A^3}\right)$. Cuando N se aproxima a su valor crítico, p se hace mucho mayor que $m_e \cdot c$. En este caso se dice que los electrones son **ultra-relativistas**.

En forma equivalente, el sistema puede estabilizarse como enana blanca siempre que la masa estelar $M \approx N \cdot (A \cdot M_n)$ no supere el valor umbral:

$M_{Ch} \approx N_{cr} \cdot (A \cdot M_n) = M_0 \cdot \left(\frac{Z}{A}\right)^2$ Esta es la denominada **masa de Chadrasekhar**.

Como M_0 se estima en 10^{31} kg, o sea unas 5 veces la masa del Sol, cuando $\frac{Z}{A} \approx \frac{1}{2}$ resulta $M_{Ch} \approx 0.25 \cdot M_0$ o sea 1.25 veces la masa del Sol. Para una esfera de hierro se obtiene $M_{Ch} \approx \left(\frac{26}{56}\right)^2 \cdot M_0$, o sea 1.2 veces la masa solar. El modelo más exacto, debido a Chandrasekhar, da para una esfera de plasma formada a partir de átomos de hierro un valor umbral de 1.44 veces la masa solar.

A partir de la ecuación para el mínimo de energía se puede hallar:

$$c \cdot p = \frac{m_e \cdot c^2 \cdot \left(\frac{N}{N_{cr}}\right)^{\frac{2}{3}}}{\sqrt{1 - \left(\frac{N}{N_{cr}}\right)^{\frac{4}{3}}}}$$

Esto a su vez junto con $R \approx \frac{\hbar \cdot c \cdot (N \cdot Z)^{\frac{1}{3}}}{c \cdot p}$ que se obtiene de los principios de Heisenberg y Pauli combinados, permite deducir una ecuación para el radio del sistema en función de su masa M :

$$R \approx \frac{\hbar}{m_e \cdot c} \cdot \left(\frac{2}{\alpha_G}\right)^{\frac{1}{2}} \cdot \left(\frac{M_0}{M}\right)^{\frac{1}{3}} \cdot \left(\frac{Z}{A}\right)^{\frac{5}{3}} \cdot \sqrt{1-\left(\frac{M}{M_{Ch}}\right)^{\frac{4}{3}}} \qquad M_{Ch} = M_0 \cdot \left(\frac{Z}{A}\right)^2$$

Si N es pequeño respecto de N_{cr}, entonces, en km: $R \approx 12000 \cdot \left(\frac{Z}{A}\right)^{\frac{5}{3}} \cdot \left(\frac{masa.solar}{M}\right)^{\frac{1}{3}}$

Cuando M es igual a la masa solar y $\frac{Z}{A} \approx \frac{1}{2}$, resulta un radio de 3.500 km.

Además, se obtiene para la energía de la enana blanca:

$$E_{\oplus,\min}(p) = \left(m_e \cdot c^2 \cdot \left(\frac{2}{\alpha_G}\right)^{\frac{3}{2}} \cdot \left(\frac{Z}{A}\right)^3\right) \cdot \frac{N}{N_{cr}} \cdot \sqrt{1-\left(\frac{N}{N_{cr}}\right)^{\frac{4}{3}}}$$

Conviene escribir en este caso $N_{cr} = \left(\frac{2}{\alpha_G}\right)^{\frac{3}{2}} \cdot \left(\frac{Z^2}{A^3}\right)$, siendo $\alpha_G = \frac{G \cdot M_n^2}{c \cdot \hbar}$ un número sin dimensiones que involucra la masa en reposo del neutrón y las tres constantes universales c, $\hbar$ y G.

Si una enana blanca cuya masa es inferior al valor crítico M_{Ch} incorpora una masa adicional tal que la masa resultante se aproxima al valor de M_{Ch}, las reacciones nucleares de fusión se reanudan en el centro de la estrella y se puede producir una explosión termonuclear gigantesca (del orden de los 10^{44} Julios) capaz de destruir la enana y expulsar violentamente el material estelar hacia el espacio. Esta catástrofe se conoce como una **Supernova Tipo I**.

12.2.2 Estrellas de neutrones

Cuando la masa de la estrella es mayor que M_{Ch}, la presión debida a los electrones no es suficiente para frenar el colapso gravitacional. Continúa la contracción hasta que se hace energéticamente favorable la reacción de captura de un electrón por un protón para producir un neutrón y un neutrino. Entonces desaparecen los electrones y protones, siendo sustituidos por neutrones. Los neutrinos son radiados al espacio exterior.

Si la masa inicial no es mucho mayor que el límite de Chandrasekhar, el sistema se estabiliza para masa del orden de los 10^{14} o 10^{15} g/cm^3, y un radio del orden de los 10 km. El campo gravitatorio confina los neutrones a distancias mutuas del orden de las dimensiones nucleares. Se conocen como **estrellas de neutrones** y se comportan como si fueran un único núcleo gigantesco.

Unos pocos centímetros cúbicos de una estrella de neutrones pesan lo mismo que el resultado de sumar la totalidad de los pesos de los habitantes de nuestro planeta en la actualidad.

Es posible construir un modelo matemático relativista, muy simplificado, para una estrella de neutrones, análogo al modelo construido para las enanas blancas.

Supongamos que la estrella está formada por N neutrones de masa en reposo M_n, y que la energía del sistema se puede reducir a un término asociado con la agitación neutrónica y un término debido a la interacción gravitatoria entre los neutrones:

$$E_{\oplus} = N \cdot \sqrt{M_n^2 \cdot c^4 + c^2 \cdot p^2} - \frac{1}{2} \cdot N^2 \cdot G \cdot \frac{(M_n)^2}{R}$$

Como las partículas no tienen carga eléctrica, el término electroestático es nulo.
La energía por neutrón se puede poner de la forma:

$$e_n(p) = \frac{E_\oplus}{N} = \sqrt{M_n^2 \cdot c^4 + c^2 \cdot p^2} - \frac{1}{2} \cdot N \cdot \frac{\alpha_G \cdot \hbar \cdot c}{R}$$

La combinación de los principios de incertidumbre y de exclusión se escribe ahora así: $\frac{R}{(N)^{\frac{1}{3}}} \cdot p \approx \hbar$ Despejando el radio del sistema e introduciéndolo en la fórmula para la energía por neutrón se obtiene: $e_n(p) = \sqrt{M_n^2 \cdot c^4 + c^2 \cdot p^2} - \left(\frac{N}{N_0}\right)^{\frac{2}{3}} \cdot c \cdot p$

Esta función presenta un mínimo cuando: $\frac{c \cdot p}{\sqrt{M_n^2 \cdot c^4 + c^2 \cdot p^2}} = \left(\frac{N}{N_0}\right)^{\frac{2}{3}}$

En esta fórmula $N_0 = \left(\frac{2}{\alpha_G}\right)^{\frac{2}{3}}$, igual que en el modelo de enana blanca. Vale 6×10^{57}.

Para que la estrella de neutrones pueda estar en un estado estable, el número de neutrones debe ser inferior a 6×10^{57}. En ese caso: $c \cdot p = \frac{M_n \cdot c^2 \cdot \left(\frac{N}{N_0}\right)^{\frac{2}{3}}}{\sqrt{1 - \left(\frac{N}{N_0}\right)^{\frac{4}{3}}}}$

Si se introduce la suma de las masas en reposo de los neutrones $M_0 = M_n \cdot N_0$, la energía del estado estable se puede estimar entonces mediante la fórmula:

$$E_{\oplus,\min} = \left(M_0 \cdot c^2\right) \cdot \frac{N}{N_0} \sqrt{1 - \left(\frac{N}{N_0}\right)^{\frac{4}{3}}} \qquad N_0 = \left(\frac{2}{\alpha_G}\right)^{\frac{2}{3}} \qquad \alpha_G = \frac{G \cdot M_n^2}{c \cdot \hbar}$$

Al número crítico de neutrones le corresponde una **masa crítica** $M_0 = N_0 \cdot M_n$ que la estrella no puede superar para poder adoptar un estado estable. Como vimos en el modelo de enana blanca, M_0 se estima en 10^{31} kg, o sea unas 5 veces la masa del Sol.
El radio de la estrella de neutrones se puede estimar mediante la fórmula:

$$R \approx \hbar \cdot \frac{(N)^{\frac{1}{3}}}{p} \approx N_0^{\frac{1}{3}} \cdot \left(\frac{\hbar}{M_n \cdot c}\right) \cdot \left(\frac{N_0}{N}\right)^{\frac{1}{3}} \cdot \sqrt{1 - \left(\frac{N}{N_0}\right)^{\frac{4}{3}}}$$

El radio característico $R_0 \approx N_0^{\frac{1}{3}} \cdot \left(\frac{\hbar}{M_n \cdot c}\right)$ es de unos 4 km.

La masa de la estrella de neutrones se puede estimar mediante la fórmula relativista:

$$\frac{E_{\oplus,\min}}{c^2} = M_0 \cdot \frac{N}{N_0} \cdot \sqrt{1 - \left(\frac{N}{N_0}\right)^{\frac{4}{3}}}$$

En función de $\frac{N}{N_0}$ la masa de una estrella de neutrones presenta un máximo igual a, aproximadamente, 2.2 veces la masa del Sol, cuando $\frac{N}{N_0} = \frac{3}{5}$. En este caso el radio vale aproximadamente 3 km.

Cuando una estrella evoluciona en forma regular y da origen a una estrella de neutrones, se contrae desde un radio del orden de los 10^5 km hasta un radio del orden de los 10 km. Cabe esperar que el momento angular $L \approx I \cdot \omega$ se conserve durante el proceso[236]. Como el momento de inercia I de la distribución de masa en torno al eje de rotación disminuye proporcionalmente a R^2 siendo R el radio de la masa en rotación, la velocidad angular ω de rotación de la estrella debe aumentar en consecuencia. Si R_0 y ω_0 representan, respectivamente, el radio y la velocidad angular de la estrella joven, mientras que R_{en} y ω_{en} representan las correspondientes variable para el estadio de estrella de neutrones, de la conservación del momento angular se desprende: $R_0^2 \cdot \omega_0 \approx R_{en}{}^2 \cdot \omega_{en}$

En la periferia de una estrella de neutrones se supone que existe una corteza de centenares de metros formada por nucleones inmersos en una masa de electrones. Las elevadas velocidades de rotación generan una estructura de campo magnético dipolar cuyo eje en general no coincide con el eje de rotación. Se supone también que la corteza eléctricamente cargada de la estrella se encuentra rodeada de una delgada fotósfera (con espesores del orden de los decímetros) que se continúa hacia el espacio exterior por una poderosa magnetósfera cuyo espesor es del orden de los 1000 km.

Desde los polos magnéticos se emiten rayos X y rayos gamma, mientras que desde la magnetósfera se emiten ondas de radio. Un observador puede medir las emisiones cuando uno de los polos magnéticos se encuentra entre el observador y el eje de rotación de la estrella: en consecuencia la detección de las emisiones es periódica, con períodos comprendidos entre algunos segundos y milésimas de segundo, dependiendo de la velocidad de rotación de la estrella de neutrones. Se observa así lo que se denomina **púlsar**.

Si la masa de la estrella considerada es más grande que 5 veces la masa del Sol, la contracción gravitatoria no puede ser contrarrestada por la agitación neutrónica y no se puede formar una estrella de neutrones sin una catástrofe previa. Se supone que la evolución de la estrella desemboca en una gigantesca explosión, denominada **Supernova Tipo II**.

En las estrellas masivas, sucesivas reacciones de fusión nuclear, precedidas de contracciones gravitatorias y elevaciones de temperatura en el centro de la masa, van involucrando elementos de número de masa cada vez mayor, hasta llegar al grupo del hierro. Llegado este punto las reacciones de fusión no pueden continuar, debido a la estabilidad de este núcleo, cuya energía de enlace es máxima. Como consecuencia de este proceso la estrella presenta una estructura de cáscaras concéntricas formadas por elementos químicos cuyo número atómico disminuye del centro a la periferia, rodeando a una esfera central de hierro: silicio, calcio, magnesio, neón, oxígeno, carbono, helio e

[236] Aunque puede variar, disminuyendo un poco, en relación con la emisión de ondas de gravitación a partir de la dinámica asociada a las irregularidades en la estructura estelar.

hidrógeno. ¡Toda una cebolla cósmica que contiene los elementos químicos necesarios para dar origen a los seres vivos!
Cuando la estrella es lo bastante masiva, la masa del corazón de hierro supera el umbral que permite la captura de electrones por los protones, se contrae aún más y los núcleos de hierro desaparecen transformándose en neutrones, emitiendo neutrinos que abandonan el sistema. Cuando la esfera central alcanza la densidad de los núcleos atómicos se supone que se frena bruscamente la contracción de la masa estelar y que se produce una onda de choque que viaja hacia la periferia provocando una explosión gigantesca. La energía de la onda de choque se obtiene de una disminución brutal y brusca de la energía gravitatoria que se estima es del orden de los 10^{44} Julios en unos segundos: una verdadera bomba gravitacional. Las capas de material estelar que rodean a la esfera central son expulsadas hacia el espacio, dando origen a una nebulosa que contiene elementos químicos precursores de la vida. En consecuencia, **la vida puede ser considerada como un emergente a partir del polvo de estrellas**. Es en este sentido en que puede interpretarse la frase, tan poética, de Hubert Reeves citada al principio del capítulo.
Como el 99% de la energía de la explosión se la llevan los neutrinos, y solamente un 1% se distribuye entre la energía cinética de los fragmentos estelares y la energía de ondas electromagnéticas, la energía luminosa en la zona visible del espectro es muy pequeña.
Las supernovas son eventos raros. En 1054 se pudo observar (por astrónomos chinos) una Supernova Tipo II, que dejó como remanente la nebulosa del Cangrejo, en la constelación de Tauro. En medio de la nebulosa se halló un pulsar, estrella de neutrones formada a partir del corazón de hierro de la estrella original.
En febrero de 1987 se pudo observar otra supernova de ese mismo tipo, en una galaxia satélite de la Vía Láctea, la gran nube de Magallanes. La detección de los neutrinos emitidos durante la explosión, cuando atravesaron nuestro planeta, aportó un elemento confirmatorio a la teoría de las supernovas. Ambos eventos, el del año 1054 y el del año 1987 se pudieron observar a simple vista[237].

13.3 Agujeros negros

Luego de la explosión, se supone que el corazón de materia estelar no dispersado puede terminar ya sea estabilizado en una estrella de neutrones, ya sea en un **agujero negro**, dependiendo de si su masa no supera o supera, respectivamente, un valor umbral relacionado con el valor crítico de 6×10^{57} del número de neutrones. Cuando ese umbral de masa es superado, la energía gravitatoria liberada por la contracción del sistema es mayor que la energía asociada a la masa de reposo de los neutrones y comienza el proceso que conduce a la formación de un **agujero negro**, dentro del cual el campo gravitatorio sería tan intenso que impediría la salida de la radiación.

Actualmente se aceptan dos tipos de agujero negro: los agujeros negros debidos al colapso gravitatorio de una estrella lo bastante masiva y los agujeros negros

[237] En cuanto a las Supernovas Tipo I, dos eventos pudieron ser observados a simple vista: uno en 1572, descripto por Tycho Brahe, y otro en 1603, descripto por Johannes Kepler.

denominados super-masivos que se manifiestan como **quásares**[238], presentan masas del orden de al menos 10^6 veces la masa del Sol, y parecen encontrarse en el centro de las galaxias (al parecer hay uno en el centro de nuestra Vía Láctea).
Otros tipos de agujero negro que han sido propuestos, como los denominados primordiales (se habrían formado por fluctuaciones cuánticas durante las primeras etapas del Universo), los mini-agujeros negros y los agujeros negros denominados de masa intermedia, no cuentan por el momento con una evidencia adecuada, ya sea basada en observaciones astronómicas, ya sea basada en los resultados de experimentos como los que se están llevando a cabo en el gigantesco acelerador de partículas, conocido como el gran chocador de hadrones (Large Hadron Collider) del CERN.

Para intentar capturar imágenes de un agujero negro, se conectaron radiotelescopios ubicados en ocho observatorios en cuatro continentes, creando el Event Horizon Telescope (EHT), equivalente a un único telescopio que tuviera casi el mismo tamaño que la Tierra.
Se tarda años en recopilar y analizar los datos obtenidos a partir de las observaciones efectuadas con el EHT, combinando métodos astronómicos y algoritmos numéricos (incluyendo herramientas de inteligencia artificial).
Recientemente los investigadores que participan en el proyecto EHT capturaron, por vez primera, cuatro imágenes de un objeto cuya masa es 6.5 billones de veces mayor que la del Sol, que se interpreta como un agujero negro. Se ubica en la constelación de Virgo, a unos 55 millones de años luz de la Tierra, cerca del centro de una galaxia conocida como M87. El conjunto de datos podría ser el más grande jamás recopilado en un trabajo científico (5 petabites). [239]

En 1798, especulando sobre las condiciones en las que la luz quedaría atrapada por el campo gravitatorio de un cuerpo, Laplace [240] publicó una deducción de lo que ahora, en el marco de la teoría generalizada de la relatividad, correspondería a lo que se denomina radio de Schwarzschild. [241]
El argumento de Laplace fue el siguiente. Para que una partícula pueda escapar del campo gravitatorio de un cuerpo esférico de masa M y radio R, partiendo de la superficie del cuerpo, debe poseer una velocidad radial v que verifique la desigualdad $\frac{1}{2}v^2 \geq \frac{G \cdot M}{R}$. Si $v^2 < \frac{2 \cdot G \cdot M}{R}$ no podrá escapar del campo de gravedad. Suponiendo que la luz se ve afectada por el campo gravitatorio como una partícula, Laplace

[238] Los quásares son fuentes de ondas de radio, sumamente intensas y situadas a distancias cosmológicas. Su enorme radiación sugiere una masa extremadamente grande, mientras que la rapidez con que varían sugiere un radio extremadamente pequeño. Por este motivo cabe suponer que son agujeros negros.

[239] Davide Castelvecchi, *Black hole pictured for first time*, Nature News 10 April 2019.

[240] P. Laplace "*Systeme du monde*", libro V, capítulo VI. Eddington llamó la atención sobre este aspecto precursor de la obra de Laplace en la página 6 de su libro "*Internal constitution of stars*", Cambridge University Press, U.K., 1926. Pero, al parecer, Laplace no fue el primero en especular sobre la posibilidad de que la luz pudiera se atrapada en un campo gravitatorio, resultando de esto la invisibilidad del cuerpo generador del campo. En 1783 John Mitchell conjeturó la existencia de estrellas oscuras para las cuales la velocidad de escape sería mayor que la velocidad de la luz.

[241] No obstante, ver las consideraciones críticas de S. Crothers en "A brief history of black holes", Progress in Physics, 2 (2006) 54-57.

sustituyó la velocidad de la luz en lugar de la velocidad de la partícula y reordenando la desigualdad dedujo que si $R < \frac{2 \cdot G \cdot M}{c^2}$ la luz quedaría atrapada por la gravedad y entonces no podríamos ver ese cuerpo: todo un agujero negro, tal como lo definió Eddington en 1935, en la cita presentada al comienzo del presente capítulo.

Ahora bien, $\frac{2 \cdot G \cdot M}{c^2} = r_S$ no es otra cosa que la fórmula para el radio de Schwarzschild. Para este valor de la coordenada radial[242] se forma una **superficie atrapada** rodeando a la distribución esférica de masa, que impide a la luz atravesarla y escapar al infinito.

Desde otro punto de vista, un agujero negro es lo que queda luego de un colapso gravitacional completo de un cuerpo. Se supone que el tiempo-espacio se encuentra tan distorsionado que ni las partículas ni la radiación pueden salir.

Un cuerpo que sea atrapado por un agujero negro pierde su identidad: dos agujeros negros formados a partir de materiales diferentes previo al colapso, no parece que puedan distinguirse uno del otro si poseen la misma masa, la misma carga eléctrica y el mismo momento angular. Estas tres propiedades influyen sobre las órbitas de los objetos que se mueven en torno o pasan cerca de un agujero negro, y así pueden en principio ser indirectamente estimadas.

El colapso de un cuerpo de masa M y radio R puede analizarse en el marco de la teoría newtoniana, considerando una masa distribuida con densidad uniforme, con una simetría esférica perfecta en torno a un centro e incapaz de desarrollar presión alguna que pudiera oponerse al efecto de la gravedad. Este podría ser un modelo de mecánica clásica, muy simplificado, de una estrella fría, que no rota.

Consideremos un elemento de masa situado a una distancia $r(t)$ del centro de la estrella. Si esta se contrae de modo tal que la densidad $\rho(t)$ permanece uniforme, entonces $r(t) = \xi \cdot R(t)$. Para cada elemento de masa ξ es un número comprendido entre 0 y 1.

Entonces si $M(r(t)) = \frac{4}{3} \cdot \pi \cdot \rho(t) \cdot r^3(t)$ es la masa encerrada en el interior de una esfera de radio $r(t)$: $\frac{d^2}{dt^2} r(t) = \xi \cdot \frac{d^2}{dt^2} R(t) = -\frac{G \cdot M(r(t))}{r^2(t)} = \frac{G \cdot \xi^3 \cdot M}{\xi^2 \cdot R^2}$

De donde el radio variable en el tiempo, verifica la ecuación diferencial $\frac{d^2}{dt^2} R = -\frac{G \cdot M}{R^2}$. Si suponemos que en el instante inicial $R(0) = R_0$ y $\frac{d}{dt} R(0) = 0$, la solución viene dada por la fórmula:

$$\sqrt{\left(1 - \frac{R}{R_0}\right) \cdot \frac{R}{R_0}} + arcsen\left(\sqrt{1 - \frac{R}{R_0}}\right) = \left(\sqrt{\frac{G \cdot M}{R_0^{\,3}}}\right) \cdot t$$

[242] Que como hemos visto en 11.2 no es igual a la distancia radial, debido a la curvatura del espacio, cosa que Laplace no podía saber dado que utilizaba un marco de mecánica newtoniana, el único disponible por aquel entonces.

A partir del análisis de esta solución se constata que el radio de la esfera disminuye entonces desde su valor inicial hasta cero en un tiempo finito, siendo ρ_0 la densidad inicial: $T_{colapso} = \frac{\pi}{2} \cdot \sqrt{\frac{R_0^3}{G \cdot M}} = \sqrt{\frac{3\pi}{16 \cdot G \cdot \rho_0}}$ La densidad final es infinita.

Si la distribución de masa no posee simetría esférica perfecta o si el material estelar es capaz de responder desarrollando un campo de presiones, el colapso hacia una singularidad caracterizada por una densidad infinita **no se produce**.

Lo visto en el capítulo 11 sugiere que las correcciones que la teoría generalizada de la relatividad introduce a la teoría newtoniana de la gravitación son proporcionales a $\frac{G \cdot M}{c^2 \cdot R} = \frac{r_S}{2 \cdot R}$. Entonces esas correcciones pueden no tenerse en cuenta en el caso de la mayoría de los objetos astronómicos, incluyendo planetas, enanas blancas, estrellas e incluso cúmulos de estrellas, porque para ellos $\frac{r_S}{R}$ es despreciable respecto de la unidad. Pero en general hay que tener en cuenta las correcciones relativistas en el caso de las estrellas de neutrones y siempre en el caso de los agujeros negros.

Un modelo relativista muy simplificado de un agujero negro, sin rotación, se puede obtener a partir de los elementos de teoría presentados en los capítulos 11 y 12, aplicados a una nube de polvo cósmico que se contrae bajo su propia atracción gravitatoria[243].
Es posible suponer que en el interior de la nube el estado del tiempo-espacio- materia se puede describir mediante la métrica estándar de Friedmann-Lemaítre-Robertson-Walker correspondiente a un espacio esférico, pero acotando la coordenada radial de modo de definir una superficie límite que permita una conexión con la materia situada en el exterior.
En el exterior la nube se supone que el tiempo-espacio se puede describir con la métrica de Schwarzschild. A medida que la contracción progresa, en este modelo la energía gravitatoria de la materia de la nube de polvo se transforma en energía cinética, permaneciendo constante la masa-energía total.

Debido a la extraordinaria densidad que cabe esperar en el interior de un agujero negro, el estado de la materia debe ser tal que para describirlo se requiere de leyes que combinan la mecánica cuántica con la relatividad generalizada en un marco de teoría unificado, todavía en desarrollo[244].
Así pues, una descripción mediante un universo esférico no puede ser adecuada desde un punto de vista astrofísico.

---

[243] R. Ruffini y J. Wheeler "*Introducing the black hole*", Physics Today, Vol. 24, 1971, pp. 30-40.

[244] Félix Cernuschi, quien dirigió mis estudios universitarios de física, llamaba **estado X** al estado de la materia en estas condiciones. Si bien estado X no es en sí mismo otra cosa que un nombre, llama convenientemente la atención sobre las diferencias que cabe esperar en relación con las propiedades conocidas de la materia, diferencias que a veces se subestiman. Cernuschi y Chandrasekhar fueron condiscípulos en uno de los cursos de relatividad generalizada que impartió Eddington, a comienzos de la década de 1930. ¡Eran los únicos alumnos en ese curso!

No obstante, en el exterior del horizonte de eventos cabe esperar que las soluciones de las ecuaciones del campo gravitatorio como la de Schwarzschild y su generalización, debida a Reissner y Nordstrom, para tener en cuenta la carga eléctrica en el caso de un agujero negro que no rota respecto de un observador alejado, o la solución de Kerr para un agujero negro en rotación, permitan una descripción adecuada desde el punto de vista astrofísico.

En general, los modelos relativistas[245] predicen que si un observador se moviera con la materia de la estrella que colapsa en el agujero negro, una vez cruzado un horizonte de eventos (la superficie atrapada) se precipitará en el interior y se vería aplastado por densidades que crecen sin límite. Luego de un intervalo finito y muy corto de tiempo propio, el observador en co-movimiento con la materia habrá completado el colapso: las dimensiones se habrán encogido hasta formar una singularidad.
Para ver cómo puede ocurrir esto, consideremos el intervalo en el caso de un movimiento a lo largo de la coordenada radial de la métrica de Schwarzschild:

$$ds^2 = c^2 \cdot \left(1 - \frac{r_S}{r}\right) \cdot dt^2 - \left(1 - \frac{r_S}{r}\right)^{-1} \cdot dr^2$$

En función del parámetro u:

$$\left(\frac{ds}{du}\right)^2 = c^2 \cdot \left(1 - \frac{r_S}{r}\right) \cdot \left(\frac{dt}{du}\right)^2 - \left(1 - \frac{r_S}{r}\right)^{-1} \cdot \left(\frac{dr}{du}\right)^2 = F^2\left(\frac{dt}{du}, r, \frac{dr}{du}\right)$$

El movimiento del cuerpo siguiendo una trayectoria $(t(u), r(u))$ en el tiempo-espacio curvado por la masa $M_\bullet$ del agujero negro y que pase por dos posiciones $(t(u_1), r(u_1))$ y $(t(u_2), r(u_2))$ dadas, deberá ser una geodésica temporaloide en ese tiempo-espacio.
En consecuencia minimizará, como vimos en 11.4.1, la integral:

$$\int_{u_1}^{u_2} F\left(\frac{dt(u)}{du}, r(u), \frac{dr(u)}{du}\right) \cdot du$$

Las ecuaciones de Euler correspondiente a este problema de cálculo de variaciones son:

$$\frac{d}{du}\left(\frac{\partial F}{\partial \frac{dt}{du}}\right) - \frac{\partial F}{\partial t} = 0 \qquad \frac{d}{du}\left(\frac{\partial F}{\partial \frac{dr}{du}}\right) - \frac{\partial F}{\partial r} = 0$$

Teniendo en cuenta la definición de F, la ecuación de Euler para la coordenada temporal se reduce a:

$$\frac{\partial F}{\partial \frac{dt}{du}} = \frac{c^2 \cdot \left(1 - \frac{r_S}{r}\right) \cdot \frac{dt}{du}}{F} = H = cte$$

Si elegimos el tiempo propio τ como parámetro, $F = c$ y en consecuencia se verifica:

[245] S. Chandrasekhar "*The mathematical theory of black holes*", Clarendon Press, Oxford, 1983.
D. Raine y E. Thomas "*Black holes: an introduction*", Imperial College Press, Londres, 2005.
G. Hooft "*Introduction to the theory of black holes*", Lectures presented at Utrecht University, 2009.

$\frac{dt}{d\tau} = \frac{H}{c \cdot \left(1 - \frac{r_S}{r}\right)}$ Sustituyendo este resultado en la expresión de $F^2 = c^2$ se deduce que $\left(\frac{dr}{d\tau}\right)^2 = H^2 - c^2 \cdot \left(1 - \frac{r_S}{r}\right)$ Si el movimiento comienza del reposo (respecto del observador inercial en el infinito) y lejos del horizonte $r = r_S$ del agujero negro, en un punto de coordenada radial $r = r_0$, con $\frac{dr}{d\tau} = 0$ en ese punto, la constante de integración queda determinada: $H^2 = c^2 \cdot \left(1 - \frac{r_S}{r_0}\right)$ Entonces $\left(\frac{dr}{d\tau}\right)^2 = c^2 \cdot \left(\frac{r_S}{r} - \frac{r_S}{r_0}\right)$

Además, si se asume que la coordenada temporal y el tiempo propio crecen al unísono de modo que $\frac{dt}{d\tau} > 0$, la constante debe ser positiva: $H = +c \cdot \sqrt{\left(1 - \frac{r_S}{r_0}\right)}$

En consecuencia: $\frac{dt}{d\tau} = \frac{H}{c \cdot \left(1 - \frac{r_S}{r}\right)} = +c \cdot \frac{\sqrt{1 - \frac{r_S}{r_0}}}{\left(1 - \frac{r_S}{r}\right)}$

Puesto que el cuerpo se dirige hacia el agujero negro, $\frac{dr}{d\tau} = -c \cdot \sqrt{\frac{r_S}{r_0}} \cdot \sqrt{\left(\frac{r_0}{r} - 1\right)}$

Separando variables e integrando entre $\tau = 0$ (cuando $r = r_0$) y τ (para un valor $r < r_0$) se obtiene: $\sqrt{(r_0 - r) \cdot r} + r_0 \cdot arcsen\left(\sqrt{1 - \frac{r}{r_0}}\right) = c \cdot \sqrt{\frac{r_S}{r_0}} \cdot \tau$

A medida que el tiempo propio aumenta, la coordenada radial disminuye aproximándose a cero, donde se encuentra la singularidad. En el límite, cuando $r = 0$, se tiene $r_0 \cdot arcsen(1) = c \cdot \sqrt{\frac{r_S}{r_0}} \cdot \tau_\bullet$. Como $arcsen(1) = \frac{\pi}{2}$, el intervalo de tiempo propio insumido en alcanzar la singularidad verifica: $\tau_\bullet = \frac{\pi}{2 \cdot c} \cdot \frac{r_0^{\frac{3}{2}}}{r_S^{\frac{1}{2}}}$

El intervalo de tiempo τ_S insumido en llegar al radio de Schwarzschild viene dado por la expresión implícita: $r_0 \cdot \sqrt{\left(1 - \frac{r_S}{r_0}\right) \cdot \frac{r_S}{r_0}} + r_0 \cdot arcsen\left(\sqrt{1 - \frac{r_S}{r_0}}\right) = c \cdot \sqrt{\frac{r_S}{r_0}} \cdot \tau_S$

Como supusimos que $\frac{r_S}{r_0}$ es despreciable, es posible desarrollar las funciones del miembro de la izquierda en potencias de $\frac{r_S}{r_0}$ con lo cual se obtiene la aproximación:

$$\tau_S \approx \frac{r_0^{\frac{3}{2}}}{c \cdot r_S^{\frac{1}{2}}} \cdot \left[\frac{\pi}{2} - \frac{2}{3} \cdot \left(\frac{r_S}{r_0} \right)^{\frac{3}{2}} \right] = \tau_\bullet - \frac{2}{3} \cdot \frac{r_S}{c}$$

Entonces el intervalo de tiempo $\tau_\bullet - \tau_S$ propio requerido para caer desde el horizonte de eventos, situado en $r = r_S$ hasta la singularidad central del agujero negro, situada en $r = 0$ se puede estimar en: $\frac{2}{3} \cdot \frac{r_S}{c}$

Así pues, un cuerpo que se deja caer desde el reposo, a partir de un valor de su coordenada radial muy grande respecto del radio de Schwarzschild de un agujero negro que no rota, emplea un intervalo finito de tiempo propio en alcanzar la singularidad situada en el origen de coordenadas. Al parecer no pasa nada extraño cuando el cuerpo, considerado como una partícula, atraviesa el horizonte del agujero negro. No obstante, si se trata de un cuerpo extenso, la curvatura del tiempo-espacio produce que las partículas que viajan por geodésicas próximas se encuentren aceleradas unas respecto de las otras, lo cual tiene como consecuencia la deformación del cuerpo extenso y su eventual destrucción[246].

A partir de las fórmulas obtenidas para $\frac{dt}{d\tau}$ y $\frac{dr}{d\tau}$ se deduce, para $r_S < r \leq r_0$:

$$t(r) = -\frac{1}{c} \cdot \sqrt{\frac{r_0}{r_S} - 1} \cdot \int_{r_0}^{r} \frac{r'^{\frac{3}{2}}}{(r' - r_S) \cdot \sqrt{r_0 - r'}} \cdot dr'$$

Cuando la coordenada radial se aproxima a $r_S = \frac{2 \cdot G \cdot M_\bullet}{c^2}$ la integral diverge a ∞.

Entonces, como la coordenada temporal corresponde al tiempo tal como lo determina un observador inercial muy alejado, respecto del cual la masa del agujero negro se halla en reposo, para este observador inercial el tiempo de aproximación al horizonte de eventos ($r = r_S$) desde $r = r_0$ es infinito.

A diferencia de la teoría basada en la física clásica, ni el apartamiento de la simetría esférica en la distribución de masa ni la presión modifica el resultado final cuando el análisis se basa en la teoría generalizada de la relatividad.

Esto ocurre por las siguientes razones. El apartamiento de la simetría esférica no importa en la medida en que el colapso se produzca en el interior de la región definida por r_S. La respuesta de presión contribuye a la masa inercial y cuando el radio de la masa que colapsa se aproxima a r_S la contribución de la presión a la masa inercial se hace comparable a la de la densidad, de modo que la presión termina por contribuir al colapso gravitatorio.

Se considera que una vez que la materia atraviesa el horizonte de eventos de un agujero negro, no lo puede abandonar: los agujeros negros no emiten ni radiación ni partículas.

[246] Ver, por ejemplo, el capítulo 6 del libro de R. Lambourne, "*Relativity, gravitation and cosmology*", Cambridge University Press, U.K. 2010.

No obstante, las partículas cargadas que pasan por las cercanías de un agujero negro y son finalmente atrapadas, alcanzan aceleraciones tan grandes al aproximarse al horizonte de eventos que emiten radiaciones X y gamma. Por este motivo se emplean detectores de rayos X ubicados en sondas espaciales con la intención de detectar agujeros negros, no a través de sus emisiones, pues los agujeros negros no emiten, sino a través de las emisiones de las partículas cargadas de la materia vecina debidas a los efectos de la interacción gravitatoria con el agujero negro.

Supongamos que un observador en co-movimiento con la materia que forma la superficie de la estrella que se colapsa, emite pulsos de ondas de radio de frecuencia f_1 (según su propio reloj) y longitud de onda λ_1 (según su propia regla) y que esas ondas son captadas por un observador que se encuentra muy alejado en relación con r_S.
Mientras que la coordenada radial R que caracteriza la posición de la superficie de la estrella sea mayor que r_S, un observador alejado recibirá la señal de radio y determinará que presenta una frecuencia que disminuye tendiendo a cero y una longitud de onda que aumenta tendiendo a infinito a medida que el radio de la estrella se aproxima al radio de Schwarzschild. Una vez que el radio R de la estrella coincide con r_S, el colapso continúa pero el observador externo no recibe señal alguna. Por tanto, si el observador se encuentra lejos, solo podrá recibir información sobre las primeras etapas del colapso. Luego de un intervalo infinito de tiempo propio de este observador externo, el colapso gravitatorio todavía no se habrá completado.
Lo que este observador vería, es que a medida que su tiempo propio tiende a infinito, se iría completando la aproximación a una configuración estática correspondiente al radio $r_S = \frac{2 \cdot G \cdot M}{c^2} \approx 3 \cdot \frac{M}{masa.solar}(km)$. Esto se conoce como "**escenario de la estrella congelada**". La relajación del radio de la estrella al radio de Schwarzschild, y por tanto, el proceso luego del cual el observador externo pierde contacto con el observador ligado a la materia que colapsa, se presenta en forma exponencial, con un tiempo característico de 10 microsegundos para una masa del orden de la del Sol.

Una consecuencia de esta doble descripción, en un caso por un observador alejado y en reposo respecto del agujero, y en el otro caso un observador en co-movimiento con la materia atraída hacia el agujero, es la imposibilidad de asignar un valor del tiempo t, en el que ocurren los eventos que determina el primer observador, para cada valor que puede tomar el tiempo propio τ, en el que ocurren de los eventos que determina el segundo observador.
Parte de la experiencia del segundo observador, acelerado respecto del primero (que a su vez puede considerarse fijo a un marco inercial), se desarrolla para valores de τ que no tienen equivalente en la experiencia del primer observador.
Parte de la evolución del universo del segundo observador debe ocurrir entonces fuera del universo del observador inercial, porque el tiempo de este último literalmente se agotó.
Como se mencionó en 8.6, este tipo de situación entra en contradicción con el principio según el cual la equivalencia entre todos los posibles sistemas de referencia para describir el mundo físico debería implicar que existe una correspondencia entre los instantes que dos observadores atribuyen al mismo evento tal que a cada instante determinado por uno de ellos le corresponda un único instante determinado por el otro

para el mismo evento. Es decir, los eventos que acontezcan en la experiencia finita de un observador también deben acontecer en la experiencia finita de otro observador: **unus mundus**, hay un solo mundo (físico).
Evidentemente las consecuencias de los modelos relativistas más aceptados para describir los agujeros negros implican que no se cumple el principio mencionado: no hay un solo mundo físico en el sentido que le hemos dado a esta expresión. [247] Recientemente se han comenzado a proponer algunos modelos alternativos, como el de S.N. Zhang y Y. Liu[248], que analizan una cáscara de materia que cae hacia un agujero negro y concluyen que el "escenario de la estrella congelada" no se presentaría en realidad: no habría singularidad ni se formaría un horizonte de eventos en la cáscara de materia durante todo el proceso.

Como las estrellas que pueden dar origen a un agujero negro por lo general rotan, un modelo más exacto requiere tener en cuenta esta rotación. En 1963 Kerr halló una generalización de la solución de Schwarzschild que permite tener en cuenta el momento angular del cuerpo[249]. La métrica de Kerr define un horizonte de eventos, pero predice la existencia de otra superficie, denominada el límite estacionario, que se ubica por fuera del horizonte de eventos, con el cual coincide en los polos donde el eje de rotación corta la superficie atrapada. (Figura 13.1)

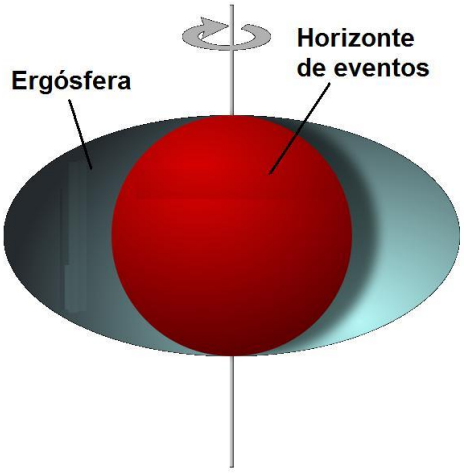


Figura 13.1 Esquema del horizonte de eventos, el límite estacionario y la región comprendida entre ambos (ergo-esfera) según la métrica de Kerr. Se indica el eje de rotación, en posición vertical.

---

[247] Este punto se discute en el artículo de R. Suárez Ántola, "Acerca del enlentecimiento de los relojes acelerados II: tiempo propio relativista y unus mundus físico", Rev. SCP, Tercera Época, 23(2): 199-226, 2018. DOI: https://doi.org/10.32480/rscp.2018-23-2.199-226

[248] S. N. Zhang y Y. Liu "Observe matter falling into a black hole", Physics Department and Center for Astrophysics, Tsinghua University, Pekín, octubre de 2007. Además, algunos autores consideran que la interpretación usual de la versión debida a Hilbert de la métrica de Schwarzschild es errónea y ponen en duda algunas de sus consecuencias y las construcciones ulteriores. Ver, por ejemplo, los artículos de L. Abrams, "Black holes: the legacy of Hilbert´s error", Can. J. Phys. 67 (1989) 919 (arXiv:gr-qc/0102055) y S. Crothers "A brief history of black holes", Progress in Physics, 2 (2006) 54-57.

[249] La métrica de Kerr fue extendida por Newman para poder describir el exterior de un agujero negro con masa, carga eléctrica y momento angular. Ver, por ejemplo S. Chandrasekhar "*The mathematical theory of black holes*", Clarendon Press, Oxford, 1983; D. Raine y E. Thomas "*Black holes: an introduction*", Imperial College Press, Londres, 2005; G. Hooft "*Introduction to the theory of black holes*", Lectures presented at Utrecht University, 2009.

La región comprendida entre el límite estacionario y el horizonte de eventos, denominada ergo-esfera a sugerencia de Wheeler y de Ruffini, se caracteriza porque al entrar en ella desde el exterior respecto del agujero negro, la coordenada temporal deja de ser una cantidad temporaloide y pasa a ser una cantidad espacialoide.

Parecería que una consecuencia de esta transmutación del tiempo sería el hacer posible una especie de convertidor de energía cósmico. Se basaría en un proceso de entrada en la ergo-esfera de un cuerpo dotado de una cierta energía positiva (asociada a su masa, es decir, energía de reposo más energía cinética) y su ruptura en dos fragmentos, uno de los cuales, juzgado por un observador muy alejado, tiene energía negativa. Asumiendo la validez del principio de conservación de la energía, el otro fragmento debe tener una energía aún mayor que la que tenía el cuerpo original al ingresar. Si este fragmento más energético escapa a lo largo de una geodésica y se aleja del agujero negro, mientras que el fragmento con energía negativa cruza el horizonte de eventos y es tragado por el agujero, entonces la energía de rotación del agujero negro habrá disminuido porque su momento angular se habrá reducido. Introdujimos una partícula con cierta energía y obtuvimos una partícula con más energía que la que ingresó y desencadenó el proceso. Este resultado, debido a Penrose[250], implica que en principio un agujero negro puede funcionar entregando energía.
Otro resultado curioso, sugerido por el análisis del colapso gravitatorio realizado por Penrose, es que el colapso puede abrir una conexión con otro universo con el cual no teníamos comunicación.
Pero mientras que el convertidor de energía cósmica es en principio comprobable por un observador externo, la apertura de una puerta a otro universo no parece serlo. Viene a cuento la parábola que Chandrasekhar aprendió en la India, siendo niño, y que aparece como última de las citas con las que comenzó este capítulo. ¿Existirá en nuestro caso el equivalente de la rana?

[250] R. Penrose "*The road to reality. A complete guide to the laws of the Universe*", Alfred Knopf, N. Y., 2005.

14. Aspectos vinculados con la lógica de la investigación científica

"Estoy completamente de acuerdo con usted sobre la importancia y el valor educativo de la metodología tanto como de la historia y la filosofía de la ciencia. Tanta gente hoy en día- aún científicos profesionales- me parece como alguien que ha visto miles de árboles pero nunca ha visto un bosque. Un conocimiento del marco histórico y filosófico suministra esa independencia de los prejuicios de su propia generación, prejuicios que sufren la mayoría de los científicos. Esta independencia, posibilitada por la penetración filosófica es-en mi opinión-la marca distintiva de un verdadero buscador de la verdad respecto de un mero artesano o un especialista."
Albert Einstein en una carta a Robert Thornton, el 7 de diciembre de 1944.

"La relación recíproca entre la epistemología y la ciencia es de un tipo notable. Son dependientes una de otra. La epistemología sin contacto con la ciencia se transforma en un esquema vacío. La ciencia sin epistemología es-en la medida en que se la pueda concebir-primitiva y embarrada. No obstante, tan pronto como el epistemólogo encuentra el sistema que busca, se ve inclinado a interpretar el contenido de pensamiento de la ciencia en términos de ese sistema y a rechazar todo lo que no se ajusta a su sistema. El científico no puede llevar tan lejos su esfuerzo por una sistematización epistemológica. Acepta agradecido el análisis de conceptos que suministra la epistemología; pero las condiciones externas que le son impuestas por los hechos experimentales, no le permiten restringirse demasiado en la construcción de su mundo conceptual por adherir a un sistema epistemológico. Por ello, el científico aparece frente al epistemólogo como un tipo de oportunista sin escrúpulos: aparece como ***realista*** *en la medida en que busca describir un mundo independiente de los actos de percepción; un* ***idealista*** *en la medida en que concibe los conceptos y las teorías como invenciones libres del espíritu humano (que no se pueden deducir en forma lógica a partir de lo empíricamente dado); como un* ***positivista*** *en la medida en que considera que sus conceptos y teorías se justifican solo en la medida en que suministran una representación lógica de las relaciones entre las experiencias sensoriales. El científico puede incluso aparecer como un* ***platónico*** *o un* ***pitagórico****, en la medida en que considera la perspectiva de la simplicidad lógica como una herramienta indispensable y efectiva para su investigación."*
A. Einstein en P. Schilpp "Albert Einstein: philosopher-scientist", The Library of Living Philosophers, Evanston, IL, 1949, pp. 683-684.

"Las ideas, al igual que los seres, no nacen adultas. Aparecen primero en medio de la confusión, entorpecidas por las mismas nociones que ellas van a invalidar, y formuladas en términos inapropiados y pronto caducos. Por este motivo las "revoluciones científicas" no alcanzan para el avance de nuestros saberes; necesariamente deben ser sucedidas por tiempos de refundación que permitan la depuración, la estabilización (provisoria) y la reformulación de teorías nuevas. Irónicamente, la física "moderna", la del siglo recién pasado, ha sufrido procesos de re-fundamentación mucho menos naturales y mucho menos eficaces que la física "clásica" del siglo XIX. Desde este punto de vista, tanto la relatividad como la mecánica cuántica han permanecido, en cuanto a la forma usual en la que se las expone, en un estado relativamente arcaico, a un siglo de su nacimiento. El contraste es chocante entre su débil evolución conceptual y terminológica y la evolución, intensa y

rápida, que sufrieron el electromagnetismo y la termodinámica. Muchas palabras y mismo conceptos que aparecen como características de estas concepciones (relatividad y cuántica) contemporáneas no son de hecho más que fósiles vivientes…"
J. Lévy-Leblond en "De la relativité à la chronogéométrie", Colloques Cargèse, "Le temps", enero de 2001.

Una vez finalizada una revisión de algunos de los antecedentes y contenidos de la teoría de la relatividad, así como de algunas de sus aplicaciones astronómicas y cosmológicas, es conveniente re-examinar la teoría, teniendo en cuenta aspectos epistemológicos, psicológicos, socio-culturales e históricos.

El valor de verdad de una teoría, visto desde un punto de vista estrictamente lógico y metodológico, es independiente de las condiciones concretas, psicológicas y socio-culturales que acompañaron su génesis.
Pero como señala José Ortega y Gasset[251], una teoría es un cuerpo de pensamientos que "nace en un alma, en un espíritu, en una conciencia, lo mismo que el fruto en el árbol…sus peculiaridades acusan ciertas tendencias específicas en el alma que la ha creado…".
Por otra parte, parecería que si una teoría de gran envergadura, como sin duda lo es la teoría de la relatividad, puede surgir y prosperar en un momento histórico determinado, ello es posible porque de un modo u otro, las personas contemporáneas, sus pensamientos, sus sentimientos, "han tomado espontáneamente determinada ruta…las ideas, cuanto más sutiles y técnicas, cuanto más remotas parezcan de los afectos humanos, son síntomas más auténticos de las variaciones profundas que se producen en el alma histórica…".
Además de las condiciones que impone el espíritu de la época en relación con las ideas de la teoría de la relatividad, es interesante prestar atención a las circunstancias de la vida personal y a las características psicológicas tanto de Einstein como de otros pioneros en la construcción de la física relativista, y cómo estas características influyeron en el desarrollo, la estructuración y la posterior aceptación generalizada de la física relativista.
Comencemos entonces por analizar en este capítulo algunos aspectos relevantes de la estructura lógica de las teorías restringida y generalizada de la relatividad, como teorías físico-matemáticas, para después considerar la cosmología desde ese mismo punto de vista. Luego, en el capítulo siguiente, centraremos nuestra atención en ciertos aspectos psicológicos y socio-culturales relacionados con el principio de relatividad, de interés desde el punto de vista epistemológico.

14.1 La estructura de las teorías físico-matemáticas

En toda teoría físico-matemática se puede distinguir tres partes indisolublemente interrelacionadas: (a) las ecuaciones propias de la teoría, (b) las reglas lógicas y las herramientas matemáticas que permiten efectuar inferencias deductivas a partir de las ecuaciones, y (c) los enunciados que establecen el significado físico de los símbolos y las palabras tales como longitud, tiempo, masa, simultaneidad, fuerza.
Es decir, una teoría física consta de lo que Carnap[252] llamaba el primer cálculo (las

[251] J. Ortega y Gasset "*El Sentido Histórico de la Teoría de Einstein*", segundo apéndice a "*El tema de nuestro tiempo*", Tecnos, Madrid, 2002.

[252] R. Carnap "*Fundamentación lógica de la física*", Sudamericana, Buenos Aires, 1969.

relaciones entre símbolos específicas de la teoría física considerada), el segundo cálculo (que establece qué transformaciones se pueden llevar a cabo a partir de las ecuaciones de la teoría sin alterar su significado) y las reglas semánticas[253] (que establecen una conexión con el lenguaje común y definen el significado físico de determinados símbolos a través de lo que Percy Bridgman[254] propuso denominar definiciones operacionales).

En una primera etapa de construcción de una teoría física predominan los procesos de inferencia no deductiva usualmente englobados en lo que se suele denominar síntesis inductiva.
Durante esta etapa, las características psicológicas de los investigadores y la influencia de la cultura dentro de la cual se lleva a cabo el trabajo, generalmente presentan una mayor incidencia respecto de lo que se observa en las etapas siguientes. Estas son la segunda etapa, de estructuración axiomática, y la tercera etapa, de deducciones cuyos resultados son contrastables experimentalmente.
Una observación cuidadosa del trabajo de las personas que construyen teorías, muestra que las tres etapas se mezclan entre sí.
En particular, los tanteos propios de la etapa inductiva se acompañan de conjeturas acerca de postulados posibles e inferencias deductivas acompañadas de comparaciones con resultados experimentales ya conocidos. No obstante, en esta primera etapa las características psicológicas condicionan el énfasis que la persona pone sobre tal o cual aspecto de los problemas examinados, orientan el esfuerzo intelectual e inciden en el clima emocional que acompaña e influye sobre los resultados del trabajo[255].

---

[253] En su acepción más general, la **semántica** comprende el estudio del significado de los símbolos, así como de las combinaciones de símbolos y las operaciones que se efectúan con símbolos. El significado de las fórmulas de una teoría física se puede abordar desde tres disciplinas: lógica, matemática y física. Desde el punto de vista lógico (semántica lógica) las fórmulas se consideran proposiciones que se suponen son verdaderas. Desde el punto de vista matemático (semántica matemática) son relaciones entre variables que toman valores numéricos, expresadas mediante igualdades (o menos frecuentemente, desigualdades). Desde el punto de vista físico (semántica física) expresan leyes de la naturaleza.

[254] P. Bridgman "*The logic of modern physics*" New York, 1929. La definición operacional de una magnitud física establece el procedimiento experimental que en cada caso permite determinar su valor, o la expresa como función de magnitudes para las que dicho procedimiento ha sido establecido. Antes de la aparición de la mecánica cuántica, los errores accidentales, inevitables en cualquier medición física, se consideraban, como fenómenos más o menos irrelevantes. Se suponía que en el límite esos errores podrían ser evitados construyendo aparatos de medida cada vez mejores: esto se relaciona con el concepto de **convergencia interna** en la medición de una magnitud física. En caso de que se disponga de varios procedimientos diferentes para determinar una misma magnitud, se exige que cumplan con el denominado requisito de **convergencia externa.** Sobre el concepto de convergencia externa, además de la obra de Bridgman, puede consultarse el libro de H. Margenau "*La naturaleza del mundo físico*", Tecnos, Madrid, 1978.

[255] Einstein tuvo una tendencia al monismo muy marcada. Siempre insistía en unificar categorías físicas diferentes interpretándolas en términos de una categoría común, presumiblemente más fundamental. Por este motivo no le satisfacía la dualidad que la relatividad restringida introduce entre el tiempo-espacio por un lado y la materia por el otro. Esta insatisfacción lo condujo a unificar la masa gravitatoria con la masa inercial primero, y a desarrollar la relatividad generalizada como una geo-metro-dinámica después. Pero no estaba satisfecho con las ecuaciones que había obtenido para el campo gravitatorio, porque del lado izquierdo aparecía el tensor métrico y sus derivadas (lado izquierdo que para Einstein era comparable con un edificio de mármol), mientras que del lado derecho aparecía el tensor de energía-impulsión (comparable, según él, con un edificio de madera de baja calidad). Esta insatisfacción lo motivó a

Pero una vez que la teoría se ha desarrollado lo suficiente, es posible deducir predicciones acerca de los resultados de experimentos u observaciones cruciales, que todavía no han sido efectuados, pero cuya ejecución resulta de suma importancia para validar o descartar la teoría. Esto a su vez conduce a considerar el requisito de falsabilidad de una teoría en el sentido establecido por Karl Popper[256].
Este requisito es particularmente aplicable a la teoría de la relatividad, tanto en su versión restringida como en su versión generalizada.
Las consecuencias de una teoría pueden ser puestas a prueba mediante observaciones cuantitativas o experimentos planeados con ese fin.
Si se encuentra una discrepancia entre las predicciones de la teoría y las observaciones, la teoría debe ser revisada: una sola instancia confirmatoria cuyos resultados se aceptan como válidos pero que resultan incompatibles con las predicciones de la teoría, basta para asegurar la falsedad de la teoría.
Pero si los resultados experimentales están de acuerdo con las predicciones de la teoría, de eso no se desprende que pueda considerársele como verdadera, sino solamente como una teoría confirmada por esos experimentos u observaciones.

A veces, a partir de una segunda teoría muy diferente, inclusive en algunos aspectos, incompatible con la primera, pueden obtenerse las mismas predicciones dentro de los límites del error experimental en un dominio de aplicabilidad común a ambas.
Este es el caso de la dinámica de la relatividad restringida y la dinámica newtoniana: todos los experimentos u observaciones que confirman la dinámica newtoniana, confirman también la dinámica relativista, pero el carácter absoluto de la simultaneidad de dos eventos en la dinámica newtoniana es lógicamente incompatible con su carácter relativo al observador en la otra dinámica.

Por otra parte es posible realizar experimentos en condiciones tales que confirman las consecuencias de la dinámica relativista pero invalidan las consecuencias de la dinámica newtoniana, y por tanto la falsean en el sentido que le daba Popper a este término.

Una situación diferente se plantea a propósitos de las **teorías paralelas**. Ahora se tienen dos o más teorías que en principio intentan dar cuenta de los mismos fenómenos, pero una de ellas no se reduce a un caso límite de otra de esas teorías, que entonces sería más general, como ocurre con la mecánica clásica y la mecánica relativista. Como vimos en 8.8, la mecánica relativista es la teoría dominante. Cuando se hace tender la velocidad de la luz a infinito se reduce a la mecánica newtoniana.
Un ejemplo de teorías paralelas respecto a la relatividad generalizada es el constituido por las teorías escalares de la gravitación que eluden la reducción de los fenómenos gravitatorios a un movimiento libre en un tiempo-espacio-materia curvado.
No obstante, casi todas las teorías escalares fracasaron en la predicción cuantitativa de

---

perseguir durante el resto de su vida, aunque sin éxito, una teoría del campo unificado. "Ante todo se requieren nuevas teorías cuando encontramos hechos nuevos que no reciben adecuada explicación en el marco de las teorías existentes. Ahora bien, tal motivo es, por así decirlo, trivial, impuesto desde afuera. Hay otra razón, más sutil y de no menor importancia. Se trata del esfuerzo por unificar y simplificar los dominios de la teoría considerada en su totalidad…" A. Einstein "Dos razones para desarrollar teorías científicas" en "*Mi visión del mundo*" (Edición de Carl Seelig, Tusquets, Barcelona, 1981).

[256] K. Popper "*La lógica de la investigación científica*", Tecnos, Madrid, 2001.

al menos uno (la rotación del perihelio de Mercurio) de los cuatro efectos cruciales que se utilizaron para confrontar las consecuencias de la teoría generalizada de la relatividad con la experiencia.
Como decía Paul Couderc[257] el proceso de geometrización propio de la teoría generalizada "no lo ha hecho tan mal en nuestra región del Universo: esperemos hasta que la rotación del perihelio de Mercurio sea explicada de alguna otra forma."

Si bien hubo que esperar hasta 1983, esa rotación fue explicada de otra manera y con el mismo grado de confirmación experimental que presenta la predicción de la teoría generalizada de la relatividad. Esto ocurrió en el marco de una teoría escalar de la gravitación que emplea el tiempo-espacio plano de la teoría restringida de la relatividad, un campo escalar para describir la gravitación y ciertas hipótesis muy simples sobre cómo el campo gravitatorio afecta localmente las reglas y los relojes. Predice igualmente bien el corrimiento al rojo, la curvatura de los rayos de luz y el retardo en los pulsos de radar, dentro de los márgenes de error alcanzables.[258]

En relación con la confirmación o invalidación experimental u observacional de las predicciones de una teoría, es importante tener presente que lo que se pone a prueba son enunciados deducidos en el marco de un sistema compuesto por las ecuaciones, las reglas de inferencia deductiva y las definiciones operacionales, y que el empleo de los instrumentos de medición y la interpretación de los resultados experimentales requiere una interpretación teórica. Así pues, cuando se intenta poner a prueba las consecuencias de una hipótesis, lo que se contrasta con la experiencia es en realidad un cuerpo de teoría completo.

Debido a este carácter holístico[259] de las teorías físicas, cuando la teoría y el experimento entran en conflicto, el conflicto puede resolverse de muchas formas diferentes.
Es posible modificar las ecuaciones, como hizo Maxwell con la formulación local de la ley de Ampère. A la densidad de corriente de conducción debida al movimiento de la carga eléctrica, le añadió la corriente de desplazamiento, para compatibilizar la formulación local de la ley de Ampère con la formulación local de la ley de conservación de la carga, para una distribución espacial continua de carga eléctrica, cuando la densidad de carga en un punto varía con el paso del tiempo.
También se pueden modificar las definiciones operacionales, o introducir una donde no

---

[257] Paul Couderc, "*The expansión of the Universe*", Mc Millan, New York, 1952.

[258] Los fundamentos de esa teoría escalar se presentan en J. Ferrari "Foundations for a theory of gravitation" Il Nuovo Cimento, vol. 78, Nº 1, 1983, pp. 53-69. Algunas aplicaciones cosmológicas se pueden hallar en J. Ferrari "Scalar gravitation and cosmology" Il Nuovo Cimento, vol. 82 B, Nº2, 1984, pp. 192-202.

[259] P. Duhem "The aim and structure of physical theory", Princeton University Press, New York, 1954. Este libro, cuya versión original, en francés, es de 1906, fue traducido al alemán en 1908 por Friederich Adler, condiscípulo de Einstein y posteriormente colega en la Universidad de Zurich. En esos años (1909-1911) al parecer los dos amigos conversaban a menudo sobre temas de filosofía de la ciencia, incluyendo los puntos de vista expuestos en las obras de Mach y de Duhem. Ver D. Howard "Einstein and Duhem", Synthese, vol. 83, 1990, pp. 363-384. En internet se encuentra el artículo de 32 páginas, también debido a D. Howard: "Einstein's philosophy of science", en la Stanford Encyclopedia of Philosophy, 2004.

las hay, como hizo Einstein en relación con el concepto de simultaneidad.
En la física clásica el concepto de simultaneidad se apoya en el postulado del transporte, que según vimos en 6.1, da por sentado que se mantienen sincronizados dos relojes que parten en el mismo instante de un punto A, de acuerdo con un reloj fijo en ese punto, y llegan simultáneamente a otro punto B, de acuerdo con el reloj fijo en ese segundo punto, pero recorren caminos diferentes para ir de A hasta B, con cualquier historia de velocidades. Einstein decidió eliminar este postulado[260] y lo sustituyó por una definición operacional de la simultaneidad para poder compatibilizar la teoría electromagnética, incompatible con la mecánica clásica, con una nueva versión de la mecánica, la mecánica relativista.

En general, para cada teoría física es posible identificar un dominio experimental dentro del cual las predicciones de la teoría están de acuerdo con los resultados de los experimentos: ese dominio se suele denominar dominio de adecuación de la teoría.
Si tenemos dos sistemas formales, ambos confirmados por los experimentos dentro de sus correspondientes dominios de adecuación, pero inconsistentes entre sí, generalmente hallamos que las inconsistencias se producen por suposiciones no operacionales, es decir, no conectadas con las mediciones u operaciones físicas en el laboratorio. Esas suposiciones no operacionales pueden ser eliminadas sin perder el grado de acuerdo con la experiencia que ya ha sido alcanzado.
De hecho, desarrollos posteriores destinados a eliminar las inconsistencias, a menudo originan una expansión extraordinaria del intervalo de concordancia con la experiencia. Eso fue lo que pasó cuando la mecánica newtoniana fue generalizada a la mecánica relativista. El dominio de adecuación de esta última teoría incluye al dominio de adecuación de la mecánica newtoniana, pero involucra además experimentos en los cuales las velocidades no son despreciables respecto de la velocidad de la luz.
Desde este punto de vista el análisis operacional de Einstein sobre el concepto de simultaneidad fue un medio para un fin, pues hizo posible integrar la mecánica de partículas con la electrodinámica clásica. El concepto de simultaneidad absoluta, fundado en el postulado del transporte, fue abandonado porque impedía la integración. Como el concepto absoluto de simultaneidad no tiene una base operacional, aunque psicológicamente es muy natural, el cambio en los fundamentos de la mecánica se pudo llevar a cabo de manera relativamente menos traumática que si hubiera sido necesario modificar una definición operacional.

A veces se puede introducir modificaciones inclusive en las reglas de inferencia deductiva, de modo que algunas de las consecuencias de un cuerpo de teoría se mantengan inalteradas.
En todo caso, lo que se confirma poniendo a prueba una consecuencia obtenida por inferencia deductiva, generalmente no es una hipótesis sino un cuerpo de teoría[261].

---

[260] Sigue habiendo polémica sobre si la definición de simultaneidad de Einstein es o no puramente convencional. Ver el artículo de A. Janis "Conventionality of simultaneity" en la Stanford Encyclopedia of Philosophy (2010).

[261] En relación con esto último, Russell (B. Russell "*The scientific outlook*", New York, 1931) ponía un ejemplo de lógica elemental, pero que constituye un análogo ilustrativo.
De las proposiciones "el material que compone el pan es harina de trigo" y "la harina de trigo es nutritiva" se desprende que "el pan es nutritivo". Supongamos ahora que alguien modifica los enunciados mencionados, de modo tal que plantea como una primera hipótesis de trabajo que "el material que compone el pan es el polvo de roca", y añade una segunda hipótesis: "el polvo de roca es nutritivo". Si se

14.2 El principio de relatividad y la teoría de la relatividad

En la teoría generalizada la materia y el tiempo-espacio se determinan entre sí, formando una totalidad indisoluble. En la teoría restringida la materia no afecta la estructura del tiempo-espacio, el cual se mantiene como algo externo que determina el modo de existencia de los fenómenos pero no se ve afectado por éstos. Esta diferencia es notable, preocupó a Einstein y lo acicateó para desarrollar la teoría generalizada y puede servir de guía con el fin de determinar y examinar posibles estructuras axiomáticas para la teoría de la relatividad.

Otra posibilidad es poner énfasis en los observadores y en la invariancia de las leyes físicas cuando se pasa de la descripción tal como la efectúa un observador a la descripción tal como la efectúa otro observador. Esto conduce a analizar ciertas simetrías inherentes al mundo físico y utilizarlas como postulados para fundamentar la teoría.

Eugen Wigner, una de las personas que en el siglo XX pensó con más profundidad sobre el concepto de simetría y sus implicaciones en las ciencias físicas, ponía énfasis en el nuevo rol que, a partir del surgimiento de la teoría de la relatividad, tuvieron las consideraciones relacionadas con la simetría en la física teórica[262].

Antes del surgimiento de la física relativista, los principios de invariancia asociados con la simetría se deducían de las leyes físicas. Después de ella, al menos parte de las leyes físicas fundamentales se dedujeron a partir de consideraciones de invariancia asociadas con simetrías en el tiempo y el espacio.

Conviene entonces partir de un postulado que introduce observadores equivalentes, cada uno con un marco espacial y un marco temporal, de modo que a menos de cambios de coordenadas admisibles, asignan una cuaterna consistente en un instante de tiempo y de tres coordenadas espaciales a cualquier suceso o evento.

Supondremos que el observador puede utilizar, al menos localmente, un espacio euclidiano para describir el ordenamiento espacial de los eventos. Para concretar, podemos considerar que emplea un sistema de coordenadas cartesiano. Las transformaciones de coordenadas admisibles, que no modifican la descripción de los eventos realizada por un mismo observador, son traslaciones y rotaciones en el espacio y traslaciones en el tiempo.

Parece razonable suponer que todos los observadores han de ser equivalentes en el sentido que todos ellos deben poder describir los mismos sucesos, cada uno desde su propia perspectiva, y que han de poder observarse entre sí. Esto excluye del ámbito de la teoría que se considera tanto el caso de los universos múltiples como algunas consideraciones especulativas sobre los agujeros negros. Por supuesto, sobre otras bases podrían no excluirse.

Suponemos ahora que para cada par de observadores y para cada suceso, existe una función que vincula las coordenadas del suceso respecto de los marcos de espacio y tiempo que le asigna un observador con las coordenadas de espacio y tiempo del mismo suceso tal como se las asigna el otro observador. Esa función se supone uno a uno y lo

---

admite la verdad de estos dos enunciados modificados, se deduce nuevamente que el pan es nutritivo. Si bien se obtiene una confirmación experimental de este último enunciado, esta confirmación no se puede asignar a las hipótesis de partida, sino solamente a la consecuencia lógica de su combinación, es decir, lo que es confirmada es una consecuencia lógica del sistema formado por las dos proposiciones y la confirmación no recae sobre las hipótesis tomadas por separado.

[262] E. Wigner "*Symmetries and reflections*", Indiana University Press, Bloomington, USA, 1967.

bastante regular como para tener derivadas parciales continuas hasta el orden que sea suficiente[263]. En símbolos: $x'^{j} = f^{j}\left(x^{0}, x^{1}, x^{2}, x^{3}\right)$ con $j = 0,1,2,3$ donde las coordenadas poseen el significado que se les atribuyó en 10.2: la primera expresa el instante y las otras tres representan las coordenadas de posición.

Al combinar las transformaciones de coordenadas admisibles para un mismo observador con las transformaciones que vinculan las coordenadas de un evento tal como son determinadas por ese observador con las coordenadas que le asigna a ese mismo evento otro observador equivalente, resulta un grupo (continuo) de transformaciones[264].

Hasta este momento hemos concretado el marco espacial y el marco temporal de cada observador. Para un observador dado tiempo y espacio aparecen separados.

Si se interrelacionan, esa interrelación solo puede aparecer en las transformaciones que conectan las coordenadas de tiempo y espacio de un mismo evento registradas por diferentes observadores. Siguiendo el camino histórico hacia la teoría de la relatividad, es en este punto donde se introduce la meditación sobre el tiempo y el planteo de una solución operacional al problema de la simultaneidad, basado en el postulado de constancia de la velocidad de la luz en los marcos de referencia inerciales, cuya existencia se admite.

Esto implica una restricción tanto sobre lo que se considera observadores equivalentes como sobre las transformaciones admisibles de coordenadas del mismo evento visto por distintos observadores.

Vía la invariancia del intervalo conduce finalmente al tiempo-espacio de Minkowski de la relatividad restringida: una variedad tetra-dimensional plana.

Luego aparece la meditación sobre la igualdad de los valores de la masa inercial y gravitatoria y la idea de igualar localmente un movimiento acelerado con un movimiento en un campo gravitatorio.

Se amplía el conjunto de observadores equivalentes y se generaliza el concepto de intervalo para corresponder a una variedad tetra-dimensional curvada por la presencia de la materia, tal que la variedad tangente en cada uno de sus puntos es un espacio de Minkowski.

Entonces, en vez de seguir el camino histórico, se puede postular desde el inicio una estructura métrica[265] (o con mayor precisión, seudo-métrica) para la variedad formada por todos los eventos posibles, variedad cuyo intervalo ds viene dado por la forma diferencial cuadrática $ds^{2} = \sum_{i,k=0}^{3} g_{i,k}\, dx^{i} dx^{k}$ caracterizada por los coeficientes tensoriales $g_{i,k}\left(x^{0}, x^{1}, x^{2}, x^{3}\right)$, tal como vimos en 9.4.3 para una variedad de n dimensiones (dimensiones que ahora restringimos a cuatro).

Vistas las cosas desde la perspectiva que brinda la simetría, lo que distingue a una teoría posible de otra es el grupo de transformaciones de coordenadas que mantiene invariante el cuadrado del intervalo entre eventos.

---

[263] Es decir, es lo que en matemática se conoce como difeomorfismo de clase C^{k}. Ver las definiciones de éste y otros conceptos relacionados con las variedades (manifolds), por ejemplo, en E. Previato (Editora) "*Dictionary of applied mathematics for engineers and scientists*", CRC press, Boca Raton, USA, 2003.

[264] Ver el diccionario de Ema Previato citado en la nota previa a ésta.

[265] Y. Ueno "Axiomatic method and theory of relativity" en L.Henkin, P. Suppes y A. Tarski (Editores) "*The axiomatic method*", North-Holland, Amsterdam, 1959, pp. 322-331.

Las transformaciones que resultan admisibles dependen a su vez de las funciones $g_{i,k}\left(x^0, x^1, x^2, x^3\right)$ que caracterizan el campo de componentes del tensor métrico. Si las leyes físicas se formulan de modo tal que su forma no cambie bajo las transformaciones admisibles, serán las mismas para todos los observadores equivalentes: precisamente por eso son equivalentes.
Tenemos entonces un **principio general de covariancia**, pero que por sí mismo no basta para construir una teoría física.
Este conjunto de suposiciones, tal como las hemos ido formulando, se podrían aplicar a la mecánica clásica tanto como a la relativista, e inclusive a otras teorías físicas posibles.
Resulta muy interesante que si a lo supuesto hasta este momento se añade un nuevo postulado que vincula dos observadores equivalentes en movimiento radial uno respecto de otro, estableciendo que si un observador detecta un evento en la línea recta que lo une a un segundo observador, entonces este último detecta el evento sobre la línea recta que lo une al primer observador, independientemente de los instantes de tiempo que cada observador asigne al evento en cuestión, entonces es posible demostrar que las transformaciones que vinculan las coordenadas de tiempo y espacio de ese evento respecto de uno y otro observador se reducen a tres casos posibles [266]: las transformaciones de Lorentz, las transformaciones de Galileo y las transformaciones K de Takeno. Estas últimas se relacionan con la expansión del universo y el principio cosmológico perfecto visto en 12.3.1.
Retomemos ahora el camino histórico para analizar otros aspectos de la estructura lógica de la física relativista.

14.3 Relatividad restringida

El principio de relatividad entendido como postulado que afirma que las leyes de la naturaleza son las mismas respectos a todos los marcos de referencia inerciales implica que una ecuación que describe una ley de la naturaleza posee la misma forma cuando se la expresa en términos de las coordenadas de tiempo y espacio correspondientes al mismo evento observado desde diferentes marcos inerciales.

14.3.1 El problema de la velocidad límite

Como vimos en 6.2, combinando este postulado con el postulado que afirma que la velocidad de la luz es la misma respecto de todos los marcos inerciales, y empleando los postulados de simetría que afirman la homogeneidad del tiempo y la homogeneidad e isotropía del espacio, se puede demostrar la invariancia de la expresión del intervalo entre dos eventos al pasar de un marco inercial a otro.
A partir de la invariancia del intervalo se pueden deducir las transformaciones de Lorentz para conectar las coordenadas de tiempo y espacio de un mismo evento respecto de dos marcos de referencia inerciales, ya sea como se hizo en 6.2 trabajando con rotaciones en un espacio tetra-dimensional, ya sea empleando una deducción más general como, por ejemplo, la presentada por Patrick Suppes [267].

---

[266] Y. Ueno y H. Takeno "On equivalent observers", Progress in Theoretical Physics, vol. 8, 1952, pp. 291-301.

[267] P. Suppes "Axioms for relativistic kinematics with and without parity", en L.Henkin, P. Suppes y A. Tarski (Editores) "*The axiomatic method*", North-Holland, Amsterdam, 1959, pp. 291-307.

Más aún, en 6.3 vimos que:
(a) Es posible prescindir del postulado de constancia de la velocidad de la luz y del intercambio de señales electromagnéticas para definir el sincronismo entre relojes ubicados en puntos diferentes de un mismo marco inercial (usando señales de naturaleza física diversa o transportando a puntos diferentes en forma cuasi-estática relojes sincronizados en un mismo punto).
(b) A partir del principio de relatividad enunciado por Poincaré, de las simetrías asumidas del tiempo y del espacio, y de la conservación de las relaciones causales entre eventos al cambiar de marco de referencia, se pueden deducir las transformaciones de Lorentz con una constante de estructura del tiempo-espacio en lugar de la velocidad de la luz.

La invalidez de la ley de adición de velocidades de Galileo resultó fundamental para la deducción de las transformaciones de Lorentz efectuada en 6.3, incluyendo el paso intermedio durante el cual se demuestra la existencia de una velocidad invariante.
Pero mientras que la constancia de la velocidad de la luz se comprueba siempre a menos de un error experimental, la no aplicabilidad de la ley clásica de adición de velocidades se desprende como conclusión definitiva a partir de los resultados experimentales obtenidos en aceleradores de partículas.
Por este motivo la deducción de las transformaciones de Lorentz (y por ende, de la relatividad de la simultaneidad) sin necesidad de postular la constancia de la velocidad de la luz, además de ser más general, posee una fundamentación experimental más firme.

En 2011 un equipo internacional de más de 200 investigadores informó los resultados de una serie de experimentos con haces de neutrinos llevados a cabo cuidadosamente en 2009, 2010 y 2011.
Los neutrinos se originaban en el acelerador del Centro Europeo para la Investigación Nuclear (CERN) y eran detectados en el Laboratorio Subterráneo de Gran Sasso (LNGS), situado a 730 km de distancia, mediante un detector especial.
Las mediciones de la velocidad de las partículas arrojan resultados que superan la velocidad de la luz en el vacío en una cantidad muy pequeña pero que al parecer, luego del correspondiente análisis del error, debe considerarse significativa[268].
Si la anomalía que se ha observado respecto de lo que cabría esperar de acuerdo con la teoría de la relatividad, se confirmara mediante experimentos independientes, y se concluyera que no se debe a un error sistemático, entonces los neutrinos podrían viajar más rápido que los fotones.
Si hay una velocidad límite, entonces debe ser un poco mayor que la velocidad de la luz en el vacío. Esto por sí solo no invalidaría las transformaciones de Lorentz tal como se dedujeron en 6.3.
No obstante, invalidaría la deducción de las transformaciones de Lorentz basada en el principio de constancia de la velocidad de la luz, y dependiendo de la interpretación que se le dé, podría obligar a un replanteo del uso de la causalidad en Física. Tampoco se puede descartar que se inicie un proceso que desemboque en una revisión radical de los fundamentos de esa super-teoria física que es la teoría restringida de la relatividad.

[268] Ver el informe "Measurement of the neutrino velocity with the OPERA detector in the CNGS beam", arXiv: 1109.4897v1, setiembre de 2011.

Posteriormente, un análisis crítico del experimento con haces de neutrinos llevado a cabo por cinco equipos de investigación independientes descubrió errores producidos durante el proceso de medición. Esto permite explicar la anomalía sin recurrir a una velocidad del neutrino superior a la de la luz en el vacío. [269]

14.3.2 Simetrías geométricas y simetrías dinámicas

Es importante destacar que el espacio de cada observador en la teoría restringida de la relatividad es un espacio euclidiano, infinito y plano [270], mientras que el tiempo se representa por un continuo unidimensional equivalente a una recta. Entonces las propiedades de simetría (homogeneidad e isotropía) poseen un carácter global, al igual que las simetrías en el tiempo-espacio de Minkowski asociadas con las transformaciones de Lorentz.

Este carácter global de las simetrías, que Wigner denominaba **simetrías geométricas**, desaparece en la teoría generalizada, siendo sustituido por simetrías locales. Como estas simetrías locales describen ciertas interacciones específicas (la gravitatoria en el caso de la teoría generalizada de la relatividad) se denominan **simetrías dinámicas**.

La condición de invariancia bajo transformaciones de Lorentz junto el requisito de que se obtenga las ecuaciones de la mecánica newtoniana de partículas como caso límite correspondiente a movimientos con velocidades muy pequeñas respecto de la velocidad de la luz, permite plantear las ecuaciones de la mecánica relativista de partículas. Las ecuaciones de la electrodinámica ya están formuladas en forma compatible con la teoría restringida de la relatividad, y el ajuste final se obtiene estableciendo las fórmulas de transformación del campo electromagnético cuando se pasa de un marco de referencia inercial a otro.

Otro enfoque posible es postular un principio de extremo que generaliza el principio de Hamilton de la mínima acción, originalmente planteado en el marco de la mecánica analítica clásica, y utilizar consideraciones de simetría para deducir la función de Lagrange (Lagrangiana) que aparece en el principio generalizado, primero para una partícula libre y después para una partícula cargada situada en un campo electromagnético[271].

La formulación relativista del principio de Hamilton se extiende sin inconvenientes a un sistema de partículas en un campo electromagnético dado. Luego se puede extender hasta reformular la teoría relativista de campos sobre la base de un principio de extremo y una Lagrangiana adecuadamente construida, cuyo dominio está incluido en el espacio de Minkowski.

---

[269] Adrian Cho "Once Again, Physicists Debunk Faster-Than-Light Neutrinos", Science Magazine, June 8, 2012

[270] Mientras no exista un campo gravitatorio, se puede utilizar como marco espacial cualquier conjunto de cuerpos cuyas distancias mutuas se mantengan constantes, es decir, que se encuentren en reposo relativo. En principio se puede pensar en añadir nuevos cuerpos, cada vez más alejados respecto del conjunto original, pero en reposo respecto de éste.

[271] Sobre este punto en particular, ver por ejemplo el artículo de R. Suárez-Ántola, "La aproximación variacional a la teoría de la relatividad: aspectos históricos y epistemológicos", Pensamiento Universitario, Nº 7, UCU, 2006, pp. 99-172. ISSN 1510-7221.
Los principios de extremo en física teórica en general, incluyendo el principio de Hamilton extendido a la física relativista, pueden estudiarse en el libro de J.L. Basdevant "*Variational principles in Physics*", Springer Verlag, New York, 2007.

14.4 Relatividad generalizada: la teoría de la gravitación y algunos de sus problemas

Los campos de fuerza presentes en el espacio absoluto de Newton y en el tiempo-espacio de Minkowski le dicen a la materia como moverse, pero los cambios de configuración de la materia, que son capaces de afectar a los campos, no tienen efecto alguno sobre esos espacios.
Por el contrario el tiempo-espacio de la teoría generalizada determina el movimiento gravitatorio de la materia y es determinado a su vez por el movimiento de la materia, de origen gravitatorio o extra-gravitatorio. El campo gravitatorio emana de toda materia, incluyendo tanto a las partículas como a las ondas electromagnéticas, y emana de sí mismo en forma de ondas de gravitación.

14.4.1 Marcos de referencia locales y coincidencias puntuales

Un nuevo principio de simetría, el principio general de covariancia se halla en la base de esta teoría generalizada.
Este principio de simetría, de naturaleza local, permite añadir observadores con marcos de referencia no inerciales al conjunto formado por observadores equivalentes.
Pero ahora, a diferencia de lo que ocurre en la teoría restringida, el marco de referencia de tiempo y de espacio de los observadores inerciales posee un carácter exclusivamente local[272].
Desde el punto de vista matemático, ese carácter local viene representado por un espacio plano de Minkowski que es tangente en cada punto de la variedad curvada formada por todos los eventos físicos posibles.
Si consideramos que la variedad curvada se encuentra embebida en un espacio euclidiano de dimensión 10 (como vimos que tiene que ser el caso a partir de las consideraciones efectuadas en 9.4.5), el espacio de Minkowski tangente se aparta progresivamente de dicha variedad a medida que nos alejamos del punto de tangencia.

Desde la perspectiva suministrada por la teoría generalizada, los puntos de la variedad tiempo-espacio no poseen realidad independiente del valor, no nulo, del tensor métrico asignado a cada punto: el tiempo-espacio es real cuando y solo cuando hay campos. Las coordenadas son simplemente rótulos para identificar eventos físicos. Lo que existe objetivamente es el tiempo-espacio-materia[273].
Entonces, para formular una teoría completa es suficiente determinar las componentes del tensor métrico en cada punto de la variedad tetra-dimensional.

---

[272] La presencia de un campo gravitatorio, en principio variable en el tiempo, modifica la distancia entre los cuerpos. Como consecuencia, a diferencia de lo que ocurre en la teoría restringida, en la teoría generalizada no se puede utilizar como marco de referencia espacial un conjunto de cuerpos en reposo relativo, tan distantes unos de otros como se quiera. La localización espacial en la teoría generalizada debe hacerse en un entorno de un cuerpo y en principio se precisaría un conjunto de cuerpos que llene la totalidad del espacio. A cada cuerpo le correspondería un reloj cuya marcha se vería influida por el campo gravitatorio local. Esta situación dista de ser trivial y a propósito de ella se plantean numerosos problemas.

[273] Si el movimiento se pudiera referir al tiempo-espacio-materia como un todo, no se habría eliminado del todo el concepto de movimiento absoluto.

Las ecuaciones del campo gravitatorio de la teoría generalizada buscan determinarlas. Como son ecuaciones a derivadas parciales, solo se pueden resolver si se introducen condiciones en la frontera adecuadas.
Eso fue lo que hizo Karl Schwarzschild cuando construyó su famosa solución de las ecuaciones de la gravitación para el caso de una masa puntual rodeada de vacío: supuso que al alejarnos de la masa, la curvatura del tiempo-espacio tiende a cero y se aproxima cada vez más a un tiempo-espacio de Minkowski.
Si la situación física no es estacionaria, es necesario especificar, además, las condiciones iniciales correspondientes.

Es importante señalar que las ecuaciones del campo no necesitan determinar el tensor métrico unívocamente, solamente deben determinar unívocamente cosas como las intersecciones de líneas de universo, las denominadas coincidencias puntuales[274].
Todas nuestras mediciones eventualmente consisten en observaciones de coincidencias puntuales, del tipo de la coincidencia de la aguja de un instrumento con una marca pre-establecida en una escala. En un mismo punto del espacio y en un mismo instante de tiempo se cruzan dos líneas de universo: la de la aguja y la de la escala.
En general, se denominan coincidencias puntuales las intersecciones de líneas de universo, es decir, los cruces entre trayectorias de cuerpos en el tiempo espacio-materia. Todas esas coincidencias puntuales son objetivas, independientes del observador. Eso se debe a que son relaciones topológicas, en el sentido siguiente: son preservadas bajo cambios arbitrarios de coordenadas, siempre que sean biunívocos y regulares, y conserven las conexiones causales entre eventos.
Cabe esperar que la formulación matemática de las leyes físicas adopte una forma invariante frente a esos cambios de coordenadas: las ecuaciones que las expresan se dice que son covariantes.

El asignar un rol fundamental a las coincidencias puntuales en la teoría generalizada fue al parecer una sugerencia de Michele Besso (amigo y ex-compañero de estudios universitarios de Einstein), realizada a Einstein durante una conversación que mantuvieron en agosto de 1913.
Einstein en un principio la rechazó, para terminar aceptándola dos años después, luego de la publicación de un trabajo de Kretschmann en el cual este autor introducía el concepto (aunque no la denominación) de coincidencia puntual: "No es necesario que las $g_{\mu,\nu}$ mismas sean determinadas unívocamente, solo deben serlo los fenómenos observables en el espacio de gravitación, es decir, el movimiento de un punto material debe serlo." [275]

A partir de ese momento, Einstein aceptó que toda nuestra experiencia física se reduce, en última instancia, a este tipo de coincidencias puntuales.

---

[274] El énfasis en las coincidencias puntuales permite eludir la presencia de un exceso de estructura matemática en la teoría generalizada, exceso que al parecer no tiene correlato físico. Este exceso se relaciona con el denominado argumento del agujero, planteado por Einstein en 1913, durante su búsqueda desesperada de las ecuaciones de la teoría generalizada. Ver el artículo de 27 páginas debido a J. Norton: "The hole argument", en The Stanford Encyclopedia of Philosophy. Revisado en 2008.

[275] Ver M. Janssen "On pots and holes", Ann. der Phys. (Leipzig), vol. 14, Supplement, pp. 58-85, 2005, y J. Norton: "The hole argument", en The Stanford Encyclopedia of Philosophy, 2008.

Reformuladas en el espíritu del Programa de Erlangen de Félix Klein, el grupo de transformaciones puntuales generales en el tiempo-espacio-materia preserva el conjunto de coincidencias puntuales, que se supone agotan el contenido esencial de las geometrías permitidas por la teoría generalizada de la relatividad. Este tipo de transformaciones preserva la diferencia entre geodésicas y no geodésicas.

Entonces es posible definir un Principio de Relatividad Generalizada: las leyes de la naturaleza son meramente enunciados acerca de coincidencias puntuales; la única forma natural de expresarlas es por lo tanto en términos de ecuaciones generales covariantes.

14.4.2 Algunas limitaciones de la teoría generalizada de la relatividad

Pese a que la teoría de la gravitación de Einstein ha sido confirmada por las observaciones y los experimentos realizados hasta la fecha, como estructura teórica presenta algunos problemas. Consideremos tres de ellos.

En primer lugar, la teoría no asigna un tensor de energía-impulsión al campo gravitatorio. Esto es debido a que reduce la gravedad a geometría del tiempo-espacio. Como consecuencia las leyes de conservación ampliadas con el campo gravitatorio se reducen a (o degeneran en) meras identidades y el campo gravitatorio no puede incorporarse en un sistema cerrado, a diferencia de otros campos asociados con la materia.

En segundo lugar, las singularidades en el tiempo espacio y los horizontes de eventos aparecen inextricablemente vinculados a la estructura de las ecuaciones de la teoría.
Al aproximarnos a las singularidades las intensidades del campo gravitatorio tienden a infinito y superan en mucho los valores que se asocian con las observaciones astronómicas y astrofísicas, y sobre los que, por tanto, se dispone de evidencia experimental.
Si se sigue adelante extrapolando la validez de las ecuaciones sin dudar, como por otra parte es usual en cosmología astrofísica, se arriba a la singularidad del Big-Bang y a una pléyade de otras singularidades, asociadas con hipotéticos agujeros negros.
Pero Einstein mismo planteó sus dudas en relación con la validez de las ecuaciones de la gravitación en condiciones tan extremas, e intentó desarrollar la teoría de modo que estas singularidades no se presentaran[276].

En tercer lugar, la teoría generalizada de la relatividad, a diferencia de otras teorías de campo, se ha mostrado bastante refractaria a ser reformulada como teoría cuántica. Esto constituye un problema por el rol central que se le asigna a la mecánica cuántica en la Física contemporánea[277].

A diferencia de la teoría restringida de la relatividad, la teoría generalizada ha sido objeto de críticas frecuentes. Han aparecido números intentos por sustituirla formulando

---

[276] A. Einstein "Physics and Reality", Journal of the Franklin Institute, vol. 221, 1936, pp. 349-382.

[277] Un resumen acerca de los puntos de vista más aceptados en la actualidad sobre las teorías de gran unificación, incluyendo la gravedad cuántica y las partículas elementales, puede hallarse en R. Adler "Gravity" y en M. Greene "Superstring theory", ambos en G. Fraser (Ed.) "*The new physics in the twenty first century*", Cambridge University Press, U.K, 2009.

teorías alternativas más o menos disímiles. Algunas de estas construcciones predicen bien los resultados de los experimentos que se consideran como experimentos cruciales para confirmar la teoría de la gravitación de Einstein[278].
Queda claro que si llega a resultar necesario sustituir la teoría restringida, entonces la teoría generalizada deberá correr la misma suerte.

14.5 Cosmología relativista

Es importante insistir en que la teoría generalizada de la relatividad es una teoría local, lo mismo que son locales las situaciones observacionales o experimentales que condujeron a mediciones confirmatorias, como la rotación del perihelio de mercurio, la curvatura de los rayos luminosos por efecto de la masa del sol, el corrimiento al rojo y el retardo de los pulsos de radar debidos a un campo gravitatorio. Todo ello referido a nuestro rincón del Universo.

Un modelo de Universo, o si se quiere, más modestamente, un modelo de Metagalaxia, se puede basar en la teoría generalizada de la relatividad pero siempre implica algo más: el principio cosmológico, las simetrías que se desprenden de este principio, el aporte de las demás ramas de la física y algunas suposiciones adicionales sobre las que muchas veces no se insiste lo suficiente.

Una de estas suposiciones se refiere a lo que se considera propiamente "Universo". Desde el punto de vista de la física relativista, todos los eventos contenidos en el cono de luz pasado de un observador, y por tanto que podrían presentar relaciones causales con el estado presente del observador, parece que deberían incluirse en su Universo. Si se le añade todos los eventos que han sufrido o pudieron haber sufrido una influencia causal de parte del observador, se tiene un conjunto de eventos máximo físicamente relacionado con el observador en cuestión: sería su Universo desde el punto de vista de las conexiones causales. Pero también se puede considerar Universo al conjunto de eventos máximo al que se le pueden aplicar las leyes físicas, necesariamente extrapoladas desde nuestro rincón en el cual observamos, experimentamos y construimos teorías. En esta acepción, el Universo de un observador podría incluir eventos que forzosamente no pueden estar causalmente relacionados con dicho observador, y podría excluir eventos que se encuentran sometidos a leyes que no se obtendrían como meras extrapolaciones de nuestras leyes de la física.

Una segunda suposición se refiere a la unicidad del Universo, que impide distinguir los aspectos esenciales de los accidentales en los fenómenos, de la forma en que esta distinción se lleva a cabo en las ciencias naturales: precisamente, comparándolos con otros fenómenos similares.
Como en cosmología esta comparación es imposible (no se dispone de un conjunto de Universos para observar y eventualmente experimentar) el método inductivo no tiene aplicación.

---

[278] Ver, por ejemplo: J. Ferrari "Foundations for a theory of gravitation" Il Nuovo Cimento, vol. 78, Nº 1, 1983, pp. 53-69; H. Yilmaz "Towards a field theory of gravitation" Il Nuovo Cimento, vol. 107, Nº 8, 1992, pp. 941-960; J. Barbour, B. Foster y N. Ó Murchadha "Relativity without relativity" Class. Quantum Grav, vol 19, 2002, pp. 3217-3248.

La imposibilidad de aplicar el método inductivo hace que la cosmología no pueda ser precisamente una ciencia natural en el mismo sentido en que lo son la física, la química o la biología.
En todo caso, como veremos luego, la cosmología física no presenta una frontera que la separe de la filosofía y de la matemática en forma tan nítida como la que parece separar a estas ciencias naturales.

Dejemos este problema de lado para considerar qué se ha hecho para eludir la imposibilidad de aplicar el método inductivo.
En su lugar se han utilizado dos métodos, que no solo no se excluyen entre sí, sino que de hecho se combinan siempre en alguna medida.

14.5.1 Dos métodos para responder al problema de la unicidad del Universo.

El primero de los dos métodos mencionados consiste en extrapolar nuestras leyes físicas en gran escala y construir modelos de universos posibles, intentando determinar el que mejor se ajusta a la evidencia empírica disponible.
El prototipo de este enfoque es la construcción de modelos cosmológicos a partir de hallar el factor de escala de la métrica Friedmann-Lemaître-Robertson-Walker utilizando las ecuaciones relativistas para el campo gravitatorio[279] y su ampliación a la teoría del Big Bang.
Se trata de una respuesta al problema de la unicidad que surge desde el centro mismo de las ciencias físicas. Entonces, como sugiere Bondi[280], el asunto de cuál es el modelo que mejor se ajusta a las observaciones pasa a ser casi un asunto secundario: la Metagalaxia es considerada como un banco de prueba de las leyes físicas. Este punto de vista es el que mejor refleja lo que se hace en la cosmología física contemporánea.
Una variante de este enfoque, que desde hace un tiempo ha adquirido cierta notoriedad, consiste en imaginar mundos posibles donde ciertas constantes de la naturaleza (tales como c, h y G) cuyo valor se conoce a partir de datos experimentales, adoptarían valores diferentes a los conocidos. Luego se intenta analizar las evoluciones de esos mundos, comparándola con la historia tal como la cuenta la teoría del Big Bang.
Las conclusiones que han desencadenado más controversia entre científicos, filósofos y teólogos, fuera del ámbito de la cosmología física propiamente dicha, son las que se relacionan con la mayor o menor verosimilitud de aparición y evolución de la vida, en uno de esos mundos y en el nuestro, según sean los valores asumidos para las constantes de la naturaleza[281].

El segundo método parte de constatar que las leyes de la física presuponen ciertas propiedades del tiempo y el espacio, que se piensa deben verificarse en gran escala.
Se toman como axiomas, se formulan matemáticamente y se procede a deducir consecuencias, añadiendo aquí y allá algunas otras suposiciones complementarias, pero

---

[279] No obstante, en 1934 Milne y Mc Crea mostraron que la mayor parte de los resultados de la cosmología relativista, incluyendo la constante cosmológica, se pueden obtener a partir de una cosmología newtoniana. Ver, por ejemplo, el capítulo 11 del libro de H. Bondi "*Cosmology*", Dover, New York, 2010, o el capítulo 3 del libro de A. Liddle "*An introduction to modern cosmology*", Wiley, New York, 2008.

[280] H. Bondi "*Cosmology*", capítulo 1, Dover, New York, 2010.

[281] Ver, por ejemplo, el libro de J. Barrow "*Las constantes de la naturaleza*", Crítica, Barcelona, 2006.

intentando mantener la coherencia lógica del sistema que se construye, e introducir el menor número de postulados que sea posible. Luego se comparan las consecuencias con las leyes obtenidas por los físicos experimentales, empleando las reglas de correspondencia entre los símbolos de la teoría y las magnitudes de la experiencia.
Se trata de una respuesta al problema de la unicidad que surge desde la frontera entre la física y la filosofía. Considera la cosmología física como una disciplina central a partir de la cual se pueden eventualmente obtener nuestras leyes físicas terrestres.
Hasta que las leyes de la física experimental permitan una confirmación o fuercen a una rectificación, ese universo al que el modelo matemático se refiere no es el real, sino uno puramente imaginario.
Y no solo se puede, sino que se debe intentar construir otros sistemas axiomáticos, otros mundos imaginarios, partiendo de otros supuestos y eventualmente, otras reglas de correspondencia.
En lugar de establecer conjeturas sobre las leyes de la naturaleza a partir de numerosas y variadas observaciones, se intenta llegar a dichas leyes por la vía de la deducción: las observaciones se limitan a confirmar o refutar el resultado obtenido.
La siguiente opinión de Eddington[282] resume muy bien esta perspectiva, cuando es llevada al extremo: "En todo el sistema de las leyes físicas no hay ninguna que no pueda ser inequívocamente deducida de consideraciones epistemológicas. Una inteligencia que no supiese nada de nuestro universo, pero que supiese cuál es el sistema intelectual mediante el cual la mente humana se interpreta a sí misma el contenido de su experiencia sensible, sería capaz de adquirir todo el conocimiento físico que nosotros hemos adquirido mediante experimentos."
El prototipo de este enfoque, actualmente muy desacreditado, es la cosmología de Milne[283].

No obstante ese descrédito, es conveniente recordar que, tanto las ciencias físicas como la cosmología física contemporáneas, al igual que todas las ciencias desde sus inicios en la antigüedad clásica, se basan en ciertos presupuestos metafísicos o epistemológicos, que generalmente no son explicitados.
Por ejemplo, el postulado de fidelidad de la naturaleza (asumir que de un momento a otro una roca no se puede transformar sin más en un chimpancé), el postulado de simplicidad (asumir que, pese a las apariencias, la estructura del mundo físico ha de ser simple, y actuar en consecuencia) o el de economía de pensamiento (trabajar con un mínimo de suposiciones independientes consistentes con una descripción mínima del objeto de estudio).

La revolución científica que se produjo al final del renacimiento y al comienzo mismo de la época moderna, no consistió tanto en una jerarquización del abordaje experimental de las ciencias en sí mismo, como en una nueva forma de abordar el diseño, la ejecución y la interpretación de los experimentos que hace de la ciencia un ejercicio controlado de la imaginación y no una mera consecuencia de la observación.

---

[282] Ver en particular la nota de A. Eddington "Physical science and philosophy", Nature, Supplement 139, 1937, pp. 1000-1001, donde explica que las leyes físico-matemáticas de la naturaleza no expresan conocimientos sobre un universo objetivo. El aspecto objetivo de nuestro conocimiento reside, según Eddington, en los detalles de eventos y sistemas particulares obtenidos a través de observaciones.

[283] H. Bondi "*Cosmology*", capítulo 10, Dover, New York, 2010.

Los experimentos de Galileo con el plano inclinado con el propósito de investigar la ley de caída de los cuerpos constituyen un ejemplo paradigmático.
La física experimental actual ha continuado, desarrollado y perfeccionado esta forma de abordar el experimento, pero no la ha modificado en lo que le es esencial[284].

14.5.2 Tiempo cósmico, espacio cósmico, marcos inerciales y modelos cosmológicos.

Dicho esto, enfoquemos la atención sobre el modelo estándar de la cosmología física actual. Su expresión matemática resulta de una selección efectuada en la clase infinita de soluciones de las ecuaciones relativistas para el campo gravitatorio, en base a suponer, no solamente que esas ecuaciones son aplicables en todas partes y en regiones de tiempo-espacio de dimensiones tan grandes como se quiera, sino admitiendo el principio cosmológico en su forma débil y el postulado de Weyl.
Pero ni el principio cosmológico ni el postulado de Weyl son parte de la teoría generalizada de la relatividad.
Son suposiciones adicionales, independientes de la teoría de la gravitación, pero con dos consecuencias cruciales para la construcción de modelos matemáticos en cosmología: la posibilidad de introducir un fluido cósmico homogéneo e isótropo para expresar el tensor de energía-impulsión, y la posibilidad de utilizar la métrica de Friedmann-Lemaître-Robertson-Walker para restringir la enorme variedad de posibles tensores métricos que podrían ser soluciones de las ecuaciones del campo gravitatorio.
Los marcos de referencia en co-movimiento con el fluido cósmico se pueden considerar como marcos inerciales locales: **la posibilidad de identificar un marco inercial queda fundamentada en un modelo cosmológico**.

Consideremos un cuerpo como la Tierra, en movimiento acelerado respecto del correspondiente marco inercial local. Supongamos que un experimento mecánico se realiza en un laboratorio fijo al suelo, se efectúa con instrumentos que poseen una cierta apreciación, y dura un intervalo de tiempo que es lo bastante pequeño como para que el cambio en la velocidad de la Tierra durante ese intervalo pueda despreciarse al interpretar los resultados de las mediciones. En ese caso, para los fines del experimento nuestro planeta puede considerarse como un sistema inercial (piénsese en un péndulo común). Si se estudian movimientos de mayor duración y en otras condiciones, esta suposición pueda ser claramente inadecuada (piénsese en un péndulo de Foucault).

Como vimos en 12.3.3, resulta posible sincronizar los relojes de todos los observadores fundamentales (en co-movimiento con el fluido cósmico) e introducir así un único tiempo cósmico y un único espacio físico tridimensional formado por todos los eventos simultáneos respecto de ese tiempo universal.
Si bien con un significado muy diferente al que tenían en física pre-relativista (puesto que resultan de una partición, posible pero muy especial, del tiempo-espacio-materia) nos encontramos finalmente con un único tiempo y un único espacio, y con la posibilidad de narrar en este marco una historia del mundo.
No obstante, las ecuaciones de la teoría relativista de la gravitación, el principio cosmológico y el postulado de Weyl por sí mismos son compatibles con muchos

[284] Sobre el significado de los experimentos de Galileo y sus continuadores, desde una perspectiva filosófica, se puede consultar el trabajo de J. Ortega y Gasset "La filosofía de la historia de Hegel y la historiología", en "*Goethe desde dentro*", Obras Completas, tomo IV, Revista de Occidente, Madrid, 1961.

modelos de Universo diferentes. Cada uno de ellos corresponde a una de las tantas soluciones de las ecuaciones del campo, dependiendo de la densidad de materia-energía y su comportamiento físico. (Figura 14.1)

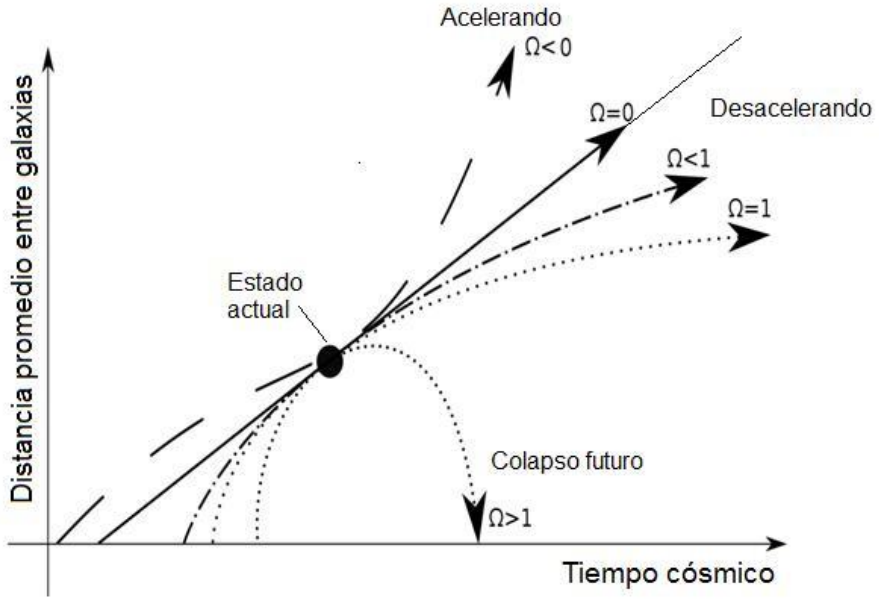


Figura 14.1 Esquema cualitativo que sugiere distintas evoluciones posibles de la distancia promedio entre las galaxias compatible con las ecuaciones del campo gravitatorio de la teoría generalizada.

La primera posibilidad en ser estudiada fue el modelo de Universo estático, que podríamos denominar de Parménides, esférico, de volumen finito, cerrado sobre sí mismo, de radio constante y eterno.
Fue propuesto por Einstein en 1917, acorde con su ideal de inteligibilidad y su convicción profunda acerca de la naturaleza ilusoria del tiempo[285].
Pero a partir de los trabajos de de Sitter (en 1917), Friedmann (en 1922 y 1924) y Lemaître (en 1927), se propusieron diversos modelos de Universos dinámicos que se expanden, que se contraen o que oscilan para siempre[286].

El tiempo en todos estos modelos cosmológicos puede ser considerado como reversible, en la medida en que las ecuaciones correspondientes se mantienen invariantes si se sustituye t por –t.
La expansión sostenida e indefinida que predice el modelo estándar, junto con la disminución de la temperatura del Universo a medida que aumenta el factor de escala, conecta el modelo con la termodinámica y sugiere algún tipo de flecha del tiempo.
Pero esa flecha es, por así decirlo, puramente geométrica, externa y no termodinámica, porque la expansión que describe el modelo es adiabática[287] (conserva la entropía del Universo).

---

[285] En una carta dirigida al hijo y a la hermana de su mejor amigo y ex-compañero de estudios, Michel Besso, con motivo de su fallecimiento, escribió, tal vez como consuelo para ellos y también para sí mismo: "Para nosotros, físicos convencidos, la distinción entre el presente, el pasado y el futuro, es una ilusión, aún cuando sea obstinada".

[286] Una presentación amena de estos primeros modelos matemáticos puede hallarse en el libro de Paul Couderc, "*The expansión of the Universe*", Mc Millan, New York, 1952.

Aun admitiendo el tipo de modelo matemático que resulta de ese enfoque del problema cosmológico, decidir en forma concluyente sobre si el modelo más adecuado es el esférico con un volumen finito aunque muy grande, o es el hiperbólico con volumen infinito y una curvatura muy pequeña, o inclusive si es el modelo plano, actualmente el favorito de la mayoría de los cosmólogos, es muy difícil debido a lo limitado del conocimiento empírico del Universo.
Cuando los criterios que deberían hacer posible la elección se confrontan con las observaciones se obtienen valores tan próximos al valor crítico correspondiente al modelo plano, que al parecer haría falta siglos o milenios de acumulación de datos para, con suerte, si el modelo más adecuado no fuera plano, obtener resultados concluyentes a favor de una de las otras dos alternativas.
Además, de acuerdo con el modelo estándar, el tiempo-espacio-materia se podría expandir tan rápido que la luz de astros lo bastante lejanos no nos podría alcanzar y entonces no la podríamos medir, no importa lo avanzados que pudieran llegar a ser nuestros instrumentos.
Esto complica todavía más la solución al problema de la finitud o infinitud del Universo.

En cuanto al problema que podríamos denominar problema de la forma del Universo, involucra tanto propiedades topológicas como métricas, y presenta dificultades aún mayores que el problema de la finitud o infinitud[288].

---

[287] En el capítulo 3 del libro de A. Liddle "*An introduction to modern cosmology*", Wiley, New York, 2008, se puede encontrar una deducción muy clara y simple de las ecuaciones de Friedmann- Lemaître involucrando como paso intermedio a la denominada ecuación del fluido cósmico. Se llega a la ecuación del fluido empleando la forma conjunta de la primera y segunda ley de la termodinámica y suponiendo un proceso adiabático y reversible para la totalidad del Universo. El modelo de Universo que se expande en estado estacionario debido a Bondi, Gold y Hoyle, se basa en la creación irreversible de materia, posee una flecha del tiempo pero no posee edad. Hace 20 años Prigogine y algunos de los integrantes de su escuela propusieron un modelo con edad y flecha del tiempo, basado en una extensión de la termodinámica de procesos irreversibles a la descripción de los procesos en el cosmos. Ver los capítulos 7 y 8 del libro de I. Prigogine e I. Stengers "*Entre el tiempo y la eternidad*", Alianza, Madrid, 1992.

[288] Este problema de la forma del Universo ha conducido a especulaciones como las denominadas **teorías de los multiversos** según las cuales nuestro Universo es uno más dentro de un conjunto de mundos que por hipótesis no pueden interactuar entre sí. Si pudieran interactuar, aún a través de una singularidad en el tiempo-espacio, formarían parte de un universo más comprehensivo. Un símil terrestre sería un conjunto de pompas de jabón, flotando alejadas unas de otras. Se distinguen dos clases de multiverso. En una primera clase se tendrían los denominados multiversos de nivel 1, en todos los cuales operarían las mismas leyes físicas, estando separados por horizontes cosmológicos. Una segunda clase, más bizarra, estaría formada por los denominados multiversos de nivel 2, en los cuales las leyes físicas variarían de un universo a otro, lo cual llevaría a pensar que ni siquiera en principio podrían ser observados. Generalmente se establece una relación entre el origen de los multiversos y el modelo inflacionario del Universo. Sobre esto pueden consultarse los artículos de divulgación de P. Steinhardt "La inflación a debate", Investigación y Ciencia, Nº 417, 2011, pp. 16-23 y de G. Ellis "¿Existe el multiverso?", Investigación y Ciencia, Nº 421, 2011, pp. 20-25. A nivel más avanzado, y abordando más en detalle algunos problemas filosóficos, se pueden consultar los trabajos de G. Ellis, U. Kirchner y W. Stoeger "Multiverses and physical cosmology", Monthly notices of the Royal Astronomical Society, vol. 347, 2004, pp. 921-936; G. Ellis "Issues in the philosophy of cosmology" en J. Butterfield y J. Earman "*Philosophy of physics*", Elsevier, N.Y. , 2006 y W. Stoeger, G. Ellis y U. Kirchner "Multiverses and cosmology: philosophical issues" arXiv:astro-ph/0407329v2, 2006.

Admitiendo el modelo plano Λ-CDM, para que este modelo resulte aplicable luego de algunos segundos después del Big Bang, el estado inicial del espacio físico debe de haber sido casi plano[289].

Para resolver este problema (entre otros, ya comentados en 12.5), es que se introdujo el proceso inflacionario del Universo. Pero con esta idea aparecen nuevas dificultades: resulta que las energías involucradas en la inflación deben haber sido tan elevadas que quedan por fuera de lo que es esperable alcanzar por la vía experimental en un laboratorio en la Tierra (se necesitaría un acelerador circular de varios años luz de diámetro), de modo que al parecer no se podría confirmar experimentalmente.

Además, si el espacio físico es plano e ilimitado, por extensa que sea la región observable del cosmos, el cociente entre el volumen de esa región accesible a la observación y el volumen del Universo siempre va a ser cero.

Esto constituye un serio problema cuando se pretende justificar el principio cosmológico débil en las observaciones que nos es dado realizar.

Así pues, vemos que la línea de demarcación entre los dos tipos de respuesta al problema cosmológico, el proveniente del centro mismo de las ciencias físicas y el proveniente de la frontera entre la ciencia y la filosofía, se ha borrado un tanto.

[289] Ver, por ejemplo, el libro de D. Ivanenko y G. Sardanashvili "*Gravitación*", capítulo 2, URSS, Moscú, 2005.

15. Perspectivas psicológicas y sociales

"…creo, con Schopenhauer que una de las razones más poderosas que impulsan al hombre hacia el arte y la ciencia es una huída de la rutina cotidiana, con su torpeza dolorosa y su yermo desconsuelo, es cortar el lazo de nuestros deseos siempre cambiantes. Es el intento de afinar las cuerdas de su existencia personal saliendo al mundo de la observación y la comprensión objetivas…El hombre procura formarse una imagen adecuada y fácilmente aprehensible del mundo, con el fin de sobreponerla a la experiencia de la realidad, sustituyéndola hasta cierto grado por ella. Esto es lo que hacen, cada uno a su modo, el pintor, el poeta, el filósofo y el investigador de la naturaleza. Hacia esta imagen y su elaboración desplazan lo principal de su vida sensible, buscando así la paz y la seguridad que no pueden encontrar en el círculo demasiado estrecho de su agitada experiencia personal."
De la conferencia sobre los principios de la investigación que Einstein desarrolló en 1927, ante los miembros de la Sociedad de Físicos Alemanes, durante el homenaje a Max Planck en el día de su 60 cumpleaños.

"La negación de la irreversibilidad del tiempo, que fue para Boltzmann una solución desesperada, llegó a ser para la mayoría de los físicos de la generación de Einstein, el símbolo mismo de lo que, para ellos, es la vocación de la física: alcanzar más allá de lo real observable una realidad temporal inteligible.
Einstein es quien mejor encarna, sin duda, el ideal que define esta nueva vocación de la física, ideal de un conocimiento que despoja nuestra concepción del mundo de lo que, a sus ojos, solo era la marca de la subjetividad humana. La ambición de ciertas prácticas místicas ha sido siempre la de escapar a las cadenas de la vida, a los tormentos y decepciones de un mundo cambiante y engañoso. En cierto sentido Einstein hizo de esta ambición la vocación misma del físico y, haciendo esto, la tradujo en términos científicos. Los místicos buscaban ***vivir*** *este mundo como una ilusión; Einstein quiso* ***demostrar*** *que es solo una ilusión, y que la verdad es un Universo transparente e inteligible, purificado de todo lo que afecta a la vida de los hombres, la memoria nostálgica o dolorosa del pasado, el temor o la esperanza del futuro."*
I. Prigogine e I. Stengers en "Entre el tiempo y la eternidad", Alianza, Madrid, 1991.

En los capítulos precedentes revisamos algunos aspectos de la historia de las ideas en torno al principio de relatividad, partiendo de la *nuova scienza* de Galileo a comienzos de la edad moderna, hasta llegar al modelo actual Λ-CDM del Universo.
Lo que examinamos desde la perspectiva que brindan las ciencias físicas y su método de investigación, puede examinarse ahora desde una perspectiva psicológica y socio-cultural.

La forma que adoptó la teoría de la relatividad, sobre todo la generalizada, se puede relacionar con ciertas características personales y presupuestos filosóficos de Einstein que revisamos para comenzar.
Luego describiremos brevemente la actividad científica y a los científicos desde las perspectivas psicológica y sociológica. Dedicamos una sección especial al reconocimiento público del aporte individual o de los equipos a la investigación. Pasamos luego a tratar la progresiva especialización y las características de los científicos que trabajan durante las etapas apolíneas con paradigmas bien consolidados, para considerar enseguida las manifestaciones de genialidad en los científicos que

realizan investigaciones en las etapas dionisíacas durante las que se reemplazan los viejos paradigmas por otros nuevos.
Finalmente describimos brevemente las características de la época moderna comparándola con la época post-moderna en la cual nos hallamos, e interpretamos la teoría de la relatividad como una teoría de transición.

La génesis y el desarrollo de los contenidos de la teoría de la relatividad, en lo que tienen de obra colectiva, se han prestado y continúan prestándose a controversia[290], en cuyo detalle no entraremos.
No obstante, más adelante en este mismo capítulo consideraremos temas que se relacionan con la mencionada controversia, a saber: las características que permiten a una persona sobresalir como un especialista genial; el efecto Mateo introducido por Robert Merton en sociología del conocimiento[291]; y lo que Simone Weil denominó el juego de ajedrez de la ciencia[292], adornado con campeonatos, premios y medallas que no han sido ajenos a los nacionalismos y a las debilidades humanas.
Pero hay un hecho que no se discute: la forma que han finalmente adoptado las teoría restringida y generalizada de la relatividad, así como sus contenidos y la manera en que se las ha dado a conocer a los físicos primero y al público en general después, se vinculan estrechamente a una sola persona, Albert Einstein, y a su visión del mundo.

15.1 Einstein y el mundo de Parménides.

Detengámonos, pues, brevemente, para señalar algunas características del que se ha constituido en el protagonista principal de esta historia.
Einstein era un físico teórico, cercano a la filosofía y alejado de la práctica de la física experimental.
Pero como todo buen físico teórico, y por ende integrador de resultados de experimentos, estaba bien enterado de lo obtenido por los físicos experimentales. Además, los trabajos tendientes a contrastar con las mediciones las predicciones de la teoría de la relatividad publicadas por Einstein, estimularon y estimulan todavía el diseño y ejecución de experimentos sofisticados, como vimos en 11.5, impulsando así el desarrollo de partes importantes de la física experimental y de la ingeniería.
Para poder formular una parte sustancial de sus ideas se requerían herramientas

---

[290] Sobre algunos aspectos que se han prestado a controversia, aunque no han pasado a ser de conocimiento público, pueden consultarse dos referencias en las que se plantea la posición minoritaria, generalmente desconocida. La primera, escrita por un distinguido físico-matemático e historiador de la ciencia inglés, es el volumen 2, "*The modern theories (1900-1926)*" del libro de E. Whittaker "*A history of the theories of aeteher and electricity*", Nelson, Londres, publicado en 1953. Allí se encuentra un capítulo sobre relatividad titulado "La relatividad de Poincaré y Lorentz" donde, haciendo referencia al famoso artículo de 1905 (sin bibliografía) titulado "Sobre la electrodinámica de los cuerpos en movimiento", se dice que Einstein publicó un artículo que exponía la teoría de la relatividad de Poincaré y Lorentz con algunos desarrollos adicionales. La otra referencia presenta argumentos que contradicen la versión oficial sobre el origen de la relatividad. Se trata del libro reciente de J. Hladik "*Comment le jeune et ambitieux Einstein s'est appropié la relativité de Poincaré*", Ellipses, París, 2004.

[291] R. Merton "The Matthew Effect in Science", Science, vol. 159, 1968, pp. 56-63. Ver también el capítulo de R. Merton y H. Zuckerman: "The Matthew Effect in Science" en R. Merton (Ed.) "The sociology of science", University of Chicago Press, Chicago, 1973 y el artículo de R. Merton "The Thomas theorem and the Matthew Effect", Social Forces, vol. 74, Nº 2, 1995, pp. 379-424.

[292] S.Weil "*Sobre la ciencia*", El cuenco de plata, Buenos Aires, 2006.

matemáticas sofisticadas, algunas de las cuales no estaban lo bastante desarrolladas en el momento en el que necesitó de ellas.
Por este motivo en el curso de su vida de investigador debió estudiar muchos conceptos matemáticos por encima de los que había internalizado durante su formación universitaria como físico profesional, y mantenerse en contacto o incluso trabajar en conjunto con matemáticos profesionales.
También en matemática estimuló la producción de trabajos originales motivados por los problemas planteados durante la formalización de las teorías restringida y generalizada.

Pese a toda esta influencia sobre las disciplinas vecinas a la física teórica y sobre la filosofía misma, Einstein, por lo que se sabe, fue la de una persona reservada, solitaria, con pocos amigos y pocos alumnos directos.
En relación con esto último, Prigogine se pregunta si no será una clave para explicar la importancia que el tema de la comunicación posee en la obra de Einstein: "su universo mental está poblado por observadores situados en marcos de referencia, en movimiento unos en relación con los otros, o sobre estrellas que poseen campos gravitacionales diferentes. Estos observadores se comunican entre ellos sus resultados mediante experiencias de emisión y recepción de señales luminosas. Es a este universo fraterno de observadores al que Einstein está ligado antes que nada. Es la objetividad incondicional que allí prevalece, lo que él está dispuesto a defender, aun al precio del abandono de la experiencia inmediata del tiempo…".[293]

Su obsesión por desvanecer el tiempo a través del pensamiento simbólico se inserta en una tradición del pensamiento de Occidente que se remonta al menos a Parménides[294].
Se encuentra en la base de su insistencia en definir la simultaneidad de dos eventos ubicados en puntos diferentes del espacio de tal forma que en última instancia el intervalo de tiempo que los separa se reduce a una consecuencia de la posibilidad de constatar sin ambigüedad la simultaneidad de dos eventos en un mismo punto y de la homogeneidad e isotropía del espacio, manifestada en el postulado de constancia de la velocidad de la luz.
También se encuentra en la concepción del continuo espacio-tiempo como una totalidad que puede ser vista como tal, por así decirlo, desde afuera.

Su marcada tendencia monista[295] lo impelía a unificar más y más.
Esto se puede constatar en su análisis del continuo espacio-tiempo de la teoría restringida.
Para fijar ideas, convengamos con Einstein en lo siguiente.

---

[293] Ilya Prigogine, en J.P.Brans, I.Stenger y P.Vincke (eds.) "*El Tiempo y el Devenir*", Gedisa, Barcelona, 1996.

[294] Un estudio detallado de la filosofía de Parménides y su relación con la búsqueda de invariantes en la ciencia puede hallarse en el libro de Karl Popper "El mundo de Parménides. Ensayos sobre la ilustración pre-socrática", capítulos 6 y 7, Paidós, Barcelona, 1999.

[295] La mente humana, cuando se ve enfrentada con la variación propia de la experiencia de la realidad, manifiesta una tendencia a buscar la identidad. Esa tendencia no puede alcanzar plenamente su objeto, puesto que la identidad pura es nada. No obstante puede progresar unificando y asimilando por grados las diversidades que encuentra en lo real. Sobre esto, ver la sección "Razón, identidad y diversidad" del capítulo sobre Psicología de la Razón que H. Delacroix escribió en el tomo 5 del "*Nuevo Tratado de Psicología*" de G. Dumas, Kapelusz, Buenos Aires, 1952.

Si un continuo físico posee las tres propiedades siguientes, entonces se lo puede considerar como un absoluto[296]:

(1) Es capaz de ejercer efectos físicos. La fuerza centrífuga y otras fuerzas inerciales se originan, tanto en la teoría restringida como en la mecánica newtoniana, en la estructura del continuo mismo[297] (el espacio absoluto en un caso, el espacio-tiempo absoluto en el otro).
(2) No recibe efectos de agentes físicos. Su estructura es rígida, no se puede influir sobre sus propiedades a través de agentes físicos.
(3) Existen propiedades geométricas que trascienden la perspectiva de cualquier observador que las determine (por ejemplo, el intervalo entre dos eventos, un invariante que trasciende las particiones del continuo tetra-dimensional en tiempo y espacio que se realizan respecto de cada marco de referencia posible).

Para Einstein, si un continuo como el de espacio-tiempo de Minkowski de la teoría restringida posee la propiedad (1), entonces no debería presentar la propiedad (2): "En primer lugar es contrario a la manera de pensar en ciencia concebir una cosa que actúa pero sobre la cual no se puede actuar."[298]
Para corregir este defecto, en la teoría generalizada se despoja de la propiedad (2) al continuo espacio-tiempo y se postula que la materia puede modificar su estructura, pero retiene las propiedades (1) y (3).

Pero Einstein no está satisfecho con este híbrido, que a una distancia prácticamente infinita de todo cuerpo material todavía exhibe estructura: la estructura plana del espacio-tiempo de Minkowski que actuaría sobre un pequeño cuerpo de prueba sin ser influido por el mismo.
"Las propiedades físicas del espacio no serían totalmente independientes, es decir, no influidas por la materia, pero en lo fundamental serían condicionadas por la materia, y eso solo en pequeña medida. Semejante situación dualista no es ni siquiera en sí misma satisfactoria."[299]
Entonces, en un espacio sin materia el tiempo-espacio no debería poseer estructura, debería perder su existencia física, desaparecer.

Einstein logró conciliar esta idea con la teoría generalizada recurriendo a un modelo

---

[296] Es interesante señalar que los puntos de vista de otros científicos eminentes sobre la importancia relativa de las propiedades que definen el continuo espacio-tiempo como un absoluto eran diferentes al de Einstein. Usaban la palabra absoluto con distinto significado. Cuando la teoría de la relatividad se hizo conocida por el gran público y comenzó a influir sobre la concepción del mundo de la época, este uso de absoluto con significados diversos produjo confusión en la interpretación de los enunciados verbales referidos al tiempo-espacio-materia. Para Max Planck, la propiedad (3) era la más importante en la definición de un continuo como absoluto, de modo que cuando apareció el tiempo-espacio-materia de la relatividad generalizada, consideró que el nuevo continuo continuaba siendo un absoluto. Para Hermann Weyl la propiedad fundamental para definir un absoluto era la (2). Como esa propiedad se perdió en el pasaje de la teoría restringida a la teoría generalizada, para Weyl el nuevo continuo no era un absoluto.

[297] H. Weyl "*Space, time matter*", Dover, New York, 1952.

[298] A. Einstein "*El significado de la relatividad*", Espasa Calpe Argentina, Buenos Aires, 1952.

[299] A. Einstein "*El significado de la relatividad*", Espasa Calpe Argentina, Buenos Aires, 1952.

cosmológico: el espacio esférico y estacionario ya mencionado en 14.5, finito y cerrado sobre sí mismo, curvado por el fluido cuyo tensor de energía impulsión (correspondiente a una distribución homogénea e isótropa) lo define y le permite existir.

15.2 El juego de ajedrez de la ciencia

En la actualidad, científicos de todos los países realizan aportes que de una u otra forma se vinculan con la teoría de la relatividad.
Pero en su origen, a fines del siglo XIX y primeros años del siglo XX, la teoría de la relatividad fue una faena típicamente europea: por así decirlo, una construcción de personas que trabajaban en el triángulo geográfico cuyos vértices se encuentran en Berlín, París y Londres.

Como toda tarea de investigación científica, fue impulsada por una pasión: la búsqueda de la verdad. Su propósito inmediato fue resolver las contradicciones entre los resultados experimentales y las predicciones de la teoría física, y entre partes de la teoría física misma, que revisamos en los capítulos 4 y 5.
La investigación científica, vista desde este punto de vista, es la continuación natural del estudio de un tema, cuando las preguntas que se nos ocurren no han sido al parecer respondidas y nos proponemos contestarlas nosotros mismos, o bien es la consecuencia natural de un hallazgo imprevisto y de la curiosidad que despierta en nosotros o en otras personas preparadas para profundizar en su comprensión.
Dependiendo del problema que se tenga entre manos, el enfoque adecuado para la investigación será ya sea experimental, ya sea teórico, o una combinación de ambos. Pero la distinción entre experimentadores y teóricos es muy marcada, y lo era ya en la época en que apareció la teoría de la relatividad.

Un experimentador debe ser paciente, perseverante para repetir el trabajo una y otra vez en caso de ser necesario, bien dispuesto para la observación concreta, para ver, oír y tocar, y dotado con habilidad manual e ingenio para diseñar, ejecutar e interpretar los resultados de sus experimentos.
Por lo general, le basta con una formación modesta en matemática.
Suele contar con la ayuda de ingenieros y técnicos de diversas especialidades, cuando es preciso diseñar y construir nuevos instrumentos, o bien modificar instrumentos conocidos.
Hasta la década de 1960, cuando entró definitivamente en la escena la así llamada "ciencia grande" y el desarrollo de la "ciencia pequeña" alcanzó un grado de sofisticación que lo hacía inconveniente, era común que muchos experimentadores diseñaran y construyeran sus propios instrumentos[300].
Un teórico, como lo fueron en su momento los protagonistas del establecimiento de las bases de la física relativista, puede bien ser algo torpe para las actividades manuales,

[300] A fines de 1969 fui aceptado para colaborar en el equipo dirigido por Antonio Frumento, por ese entonces contratado en la universidad donde estudié medicina. Frumento se construía, con ingenio y gran destreza, parte de los aparatos que utilizaba en sus investigaciones sobre propiedades eléctricas de las membranas biológicas. Una vez, para determinar las concentraciones de ciertos iones en el interior de fibras musculares esqueléticas, construyó un espectrómetro y me permitió participar. Esta actitud le resultaba bastante desconcertante a mi padre, según me lo manifestó en reiteradas oportunidades, porque le parecía una pérdida innecesaria de tiempo. Pero Frumento tenía sus motivos. Hoy sería difícil justificar esa actitud, excepto en situaciones muy especiales de la ciencia pequeña.

porque sus instrumentos de trabajo son lápiz, papel y hoy día, ordenador (en otra época, regla de cálculo y tablas numéricas). Puede no estar dotado para la observación sensorial. Pero debe estar especialmente dotado para el pensamiento abstracto y debe conocer y dominar perfectamente aquellas partes de la matemática relacionadas con su trabajo. Actualmente, además, debe comprender a fondo las posibilidades y las limitaciones de las técnicas de simulación digital, incluyendo los paquetes de software que puede emplear en su trabajo. Siempre debe estar bien familiarizado con los resultados experimentales de las disciplinas en las que trabaja.

Ya en el siglo XIX, pocos físicos se destacaron a la vez como experimentadores y como teóricos, en parte por las diferentes características psicológicas asociadas con una y otra actividad, difíciles de hallar combinadas en una sola persona, en parte por la especialización de la actividad de investigación en física, que ya se hacía notar.
El número de estos investigadores excepcionales, relativo al total, ha sido todavía menor durante el siglo XX. Esto no sorprende, dado el significativo e irreversible avance de la especialización en todas las ciencias y en las técnicas que se produjo en el siglo pasado.

Pero la investigación científica no se resume en estos aspectos tan apacibles.
Es siempre más que eso. Una de las razones que conducen a comparar la actividad científica con un juego de ajedrez es un hecho en sí mismo nada sorprendente: la competencia entre los científicos, incluso la lucha entre ellos, y algunas veces la competencia desleal.
Esta competencia se produce con el propósito de ser primeros en resolver un problema, descubrir un fenómeno o plantear una idea novedosa.
Ser los primeros tiene como consecuencia el reconocimiento de sus pares, facilidad para acceder a recursos monetarios para sí mismos y sus colaboradores, estabilidad para continuar trabajando (y viviendo), y honores gratificantes en el ámbito académico.
Si el aporte que se les atribuye es lo bastante importante, reciben un reconocimiento público, premios, cantidades de dinero significativas, tal vez fama imperecedera…

Esta dura competencia entre científicos es bien conocida entre los del oficio.
Si bien lo es menos entre el gran público, éste es consciente de algunos ejemplos de esa lucha competitiva durante la edad moderna (tal como Newton versus Leibniz) y más aún durante la post-moderna.
Cuando opera en forma adecuada, la competencia entre personas o equipos de investigación es uno de los principales mecanismos impulsores del avance en las ciencias, junto con la curiosidad acompañada por una formación adecuada y el deseo de salir a la caza de la realidad, completada por la evaluación rigurosa del trabajo del científico por otros científicos (la evaluación por pares).

No estuvo ausente en el desarrollo de la teoría de la relatividad durante sus comienzos, con la peculiaridad de que se enmarcó en la rivalidad existente entre Francia y Alemania previa a la Primera Guerra Mundial y posterior a la guerra franco-prusiana.
Como lo expuso claramente Poincaré, en 1904, al público que asistió a su conferencia[301] en el Congreso Internacional de las Artes y las Ciencias durante la exposición de Saint

[301] H. Poincaré, "Los principios de la física matemática", en Albert Einstein y otros, "*La teoría de la relatividad*", Altaya, Barcelona, 1993.

Louis, lo que estaba a la vista era, ni más ni menos, que una nueva mecánica destinada a sustituir a la clásica.
Por este motivo la competencia por la prioridad sobre las nuevas ideas fue apasionada e involucró a personas destacadas de uno y otro país, y también de otros países europeos.

La actitud fue muy diferente en Francia y en Alemania.
En parte esto fue debido a la idiosincrasia de cada uno de sus pueblos, en parte se debió a una actitud diferente de las élites científico-técnicas y políticas en Alemania y en Francia (aquí los aspectos ideológicos no estuvieron ajenos), y en parte se vinculó con la diferente circunstancia vital y diferentes características de personalidad en los protagonistas de una y otra parte.

Pudo haber influido también el equivalente en ciencia a un hecho que Picasso señaló en el caso de la innovación en las artes plásticas. En palabras de Gertrude Stein[302]:
"Uno no siempre entiende, antes de que (las obras) estén completamente terminadas, qué está ocurriendo y uno no comprende en absoluto lo que ha hecho hasta el momento en que todo está hecho. Picasso dijo una vez que aquel que creaba una cosa estaba forzado a hacerla fea. En el esfuerzo por crear la intensidad, y en la batalla para crear esta intensidad, el resultado siempre produce cierta fealdad. Los que siguen pueden hacer de esto algo bello, puesto que saben lo que están haciendo, una vez que las cosas han sido inventadas, pero el inventor, como no sabe lo que va a inventar, inevitablemente lo que inventa debe tener su fealdad."

En todo caso la tendencia imperante en los países europeos, a fines del siglo XIX y comienzos del siglo XX, era minimizar u ocultar los aspectos reprobables de la competencia entre países y enfatizar sus aspectos positivos sobre el avance del conocimiento[303].
Tal vez esto explique la forma algo extraña en la que el Profesor Whittaker aborda la historia de la teoría de la relatividad en su libro sobre el desarrollo de las ideas relacionadas con el electromagnetismo, citado a comienzos de este capítulo.
Posiblemente, otra consecuencia fue el contenido de la justificación del Premio Nobel que se le otorgó a Einstein en 1921.
Allí no se hizo mención a la teoría de la relatividad, ni a la restringida ni a la generalizada[304], y sí a sus significativos aportes a la comprensión del efecto fotoeléctrico, basados en una extensión de la idea del quantum de su mentor, Max Planck.
Pero luego se impuso en forma definitiva el punto de vista actual sobre el asunto de la paternidad de la teoría de la relatividad, y los cuestionamientos desaparecieron o se

---

[302] G. Stein "Picasso", Batsford, Londres, 1948, p.9, citada por R. Alpher "Una visión retrospectiva del modelo cosmológico del Big Bang" en J. Gonzalo, J. Sánchez y M. Alario (Eds.) "*Cosmología astrofísica*", Alianza, Madrid, 1995.

[303] En general se ocultaban las imperfecciones hasta un punto que un joven y falto de experiencia Henri Bergson, se consideraba peor desde el punto de vista ético, que las personas que lo rodeaban, y se mortificaba con esos pensamientos.

[304] Pese a que Einstein ya era famoso por ellas y a que en relación con su teoría de la gravitación, la que por aquel entonces podía considerarse más discutible, se habían efectuado con la participación de Inglaterra, con resultado positivo, los experimentos confirmatorios de 1919.

mantienen circunscritos a ciertos círculos muy poco numerosos, particularmente en Francia.

15.3 El reconocimiento individual a la labor científica.

Los descubrimientos y las ideas nuevas importantes se personalizan siempre, utilizando el nombre de las personas para aludir brevemente a su descubrimiento o a lo que se cree ha sido su descubrimiento.
Aun cuando varias personas, en forma independiente, hallan arribado al mismo resultado, por lo general termina triunfando un nombre, una obra, que se cita cada vez más hasta que desaparecen las demás de la conciencia de los investigadores y pasa posteriormente a ser la única reconocida por el público.
Algunas veces, al parecer más a menudo de lo que sería de esperar, se verifica lo que en sociología de la ciencia se conoce como "Ley de Stiegler". Con cierta ironía, se resume así, en su versión fuerte: ningún descubrimiento científico se denomina de acuerdo a su verdadero descubridor[305].
Esto puede deberse a que algunos descubrimientos son prematuros[306]: en el momento en que ocurren no se los puede conectar en forma satisfactoria con el conocimiento canónico de la época. El descubrimiento, debido a Mendel, de ciertas leyes estadísticas y su interpretación, efectuada en 1865, en términos de algo que posteriormente se conoció como gen, es un caso típico. Otro caso de prematuridad fue la teoría de Polanyi, formulada entre 1914 y 1916, sobre la adsorción de gases sobre superficies sólidas, que anticipaba ciertas ideas sobre lo que serían conocidas, después de 1930, como resonancias mecánico-cuánticas. Además, este tipo de obra prematura adolece a menudo de la fealdad que Picasso sugiere en relación con la verdadera innovación en artes plásticas, lo cual contribuye a dificultar aún más su aceptación.

Cuando el descubrimiento no es prematuro, pero varias personas lo efectúan en forma más o menos simultánea, a veces las características de una de ellas le hacen sobresalir e imponer su paternidad, mientras que lo realizado por los demás es relegado a la sombra.
Puesto que los recursos no crecen a la par con el número de científicos, la obtención de resultados verdaderamente originales e importantes se ha hecho cada vez más difícil. Esta situación, en el contexto de una fuerte competencia y una especialización creciente, hace que la mayor parte de los trabajos se realicen en equipo, y que varios equipos arriben a un mismo resultado.
Pero en este caso es común que se manifieste el efecto Mateo, entre equipos o dentro de un mismo equipo.
La denominación del efecto, bien conocido en sociología del conocimiento, hace referencia a un pasaje del Nuevo Testamento según san Mateo: "A todo el que tiene, se le dará y le sobrará; pero al que no tiene, aun lo que tiene se le quitará."
Cuando se produce dentro de un mismo equipo, el efecto se manifiesta con la atribución de todo el crédito al integrante más conspicuo.

---

[305] S. Stigler "Stigler´s law of eponymy" en Th. Gieryn (Ed.) "Science and social structure", Transactions of the New York Academy of Sciences, series II, vol.39, 1980, pp. 147-157.

[306] G.Stent "Las paradojas del progreso", capítulo 5, Alhambra, Madrid, 1981.

Un ejemplo es la atribución a George Gamow de resultados relacionados con la teoría del Big Bang debidos a Ralph Alpher (que fue alumno de postgrado y colaborador de Gamow) trabajando en equipo con Robert Herman (que no lo fue), quienes predijeron el fondo de radiación cósmica fósil en 1948.
No fueron reconocidos, aun estando vivos (Gamow ya no lo estaba y por ende no podía ser premiado) cuando se decidió entregar el Premio Nóbel, en 1978, a Penzias y a Wilson, por su descubrimiento[307] de dicha radiación electromagnética en 1964.

A menudo, una vez puesto en marcha, el proceso avanza hasta un punto en el que al investigador famoso se le atribuyen resultados en cuya obtención no tuvo nada que ver.
A esto se añade lo siguiente: los comités que proponen candidatos para ser premiados tienden a premiar a científicos de su misma corriente de ideas, lo cual termina generando una acumulación de premios en las mismas escuelas o en escuelas muy similares. Esto, a su vez, les da una ventaja para obtener fondos tanto públicos como privados[308].
Aun así, los resultados del análisis filosófico y sociológico sugieren que en relación con los trabajos de investigación en el marco de paradigmas bien establecidos, todo redunda en un mayor beneficio para la sociedad[309].

Cuando la investigación no se enmarca en un paradigma bien establecido sino que tiende a reemplazarlo, en el sentido que desde Kuhn se le da a esta palabra como vimos en 8.8 , o para decirlo de otra forma y a la manera de Nietzsche, cuando la investigación ya no es apolínea sino dionisíaca, el efecto (que sería en última instancia socialmente beneficioso) de los procesos que acabamos de considerar ya no se manifiesta y se transforman en un verdadero impedimento para el avance del conocimiento.
Ahora el sistema de evaluación por pares muestra sus deficiencias, al igual que los mecanismos de asignación de recursos para la investigación.
En particular, los investigadores jóvenes y desconocidos tienen más dificultad para ver publicadas ideas realmente innovadoras, y ese efecto se refuerza cuando el trabajo se realiza en un país de la periferia.
En este último caso no es raro que un trabajo sea rechazado con argumentos que indican que no ha sido examinado en profundidad y que las verdaderas debilidades que pueda presentar no sean detectadas, con lo cual el autor no se beneficia, como debería ser, con los comentarios del árbitro[310].

---

[307] Si se lo piensa en relación con el Big Bang, el hallazgo de Penzias y Wilson fue sin duda casual.

[308] H. Zuckerman "*The scientific elite: Nobel laureates in the United States*", Free Press, New York, 1977.

[309] M. Stevens "The role of the priority rule in science", Journal of Philosophy, vol. 100, 2003, pp. 55-79. M. Stevens "The role of Matthew effect in science", History and Philosophy of Science, Vol. 37, No. 2, 2006, pp. 159-170.

[310] En 1981 un joven alumno mío presentó un trabajo notable de relatividad restringida, como requisito para la aprobación de una asignatura de la carrera de física en una universidad donde fui docente. Contenía resultados que parecían originales, sobre la relación entre la homogeneidad del tiempo y la homogeneidad e isotropía del espacio, y en su deducción de las transformaciones de Lorentz no involucraba ni la suposición de Einstein sobre la constancia de la velocidad de la luz ni la existencia de una velocidad invariante (como supusimos en 6.3 para deducir las transformaciones de Lorentz sin recurrir al segundo postulado de Einstein). Antes de decidir si enviarlo o no a una revista científica, me pareció conveniente consultar a J.D.J., un eminente físico teórico norteamericano, que presumiblemente

15.4 Especialización, genialidad y desmesura.

El extraordinario desarrollo de la ciencia y la técnica basada en la ciencia, que se puede observar durante el siglo XIX, y más aún durante el siglo XX, conlleva la pérdida progresiva de contacto del científico y el técnico con las demás partes de estas disciplinas. Este es un efecto natural de la especialización creciente, que reduce cada vez más el ámbito que la mayor parte de los científicos y los técnicos pueden conocer a fondo y donde pueden efectuar aportes originales.

A comienzos de la década de 1930, en su libro más famoso[311], Ortega y Gasset advertía sobre las consecuencias de este proceso:
"La especialización comienza, precisamente, en un tiempo que llama hombre civilizado al hombre "enciclopédico". El siglo XIX inicia sus destinos bajo la dirección de criaturas que viven enciclopédicamente, aunque su producción tenga ya un carácter de especialismo. En la generación sub-siguiente, la ecuación se ha desplazado, y la especialidad empieza a desalojar dentro de cada hombre de ciencia a la cultura integral. Cuando en 1890 una tercera generación toma el mando intelectual de Europa, nos encontramos con un tipo de científico sin ejemplo en la historia. Es un hombre que, de todo lo que hay que saber para ser un personaje discreto, conoce solo una ciencia determinada, y aún de esa ciencia solo conoce bien la pequeña porción en que él es activo investigador. Llega a proclamar como una virtud el no enterarse de cuanto quede fuera del angosto paisaje que especialmente cultiva, y llama diletantismo a la curiosidad por el conjunto del saber.
El caso es que, recluido en la estrechez de su campo visual, consigue, en efecto, descubrir nuevos hechos y hacer avanzar su ciencia, que el apenas conoce, y con ella la enciclopedia del pensamiento, que concienzudamente desconoce. … La razón de ello está en lo que es, a la par, ventaja mayor y peligro máximo de la ciencia nueva y de toda la civilización que esta dirige y representa: la mecanización. …Así la mayor parte de los

---

debería estar muy versado en el asunto que nos ocupaba. En su contestación, que se demoró bastante por sus múltiples ocupaciones, manifiesta que no le parece que el trabajo incluya contenidos realmente nuevos, ya que siempre es un asunto de gustos lo que uno asume y lo que uno deduce. En seguida me hace notar que si en la ecuación 42 el autor hubiera efectuado el segundo postulado de constancia de la velocidad de la luz, se obtendría la ecuación 53. Ahora bien, ¡el objetivo del trabajo era precisamente llegar a la ecuación 53 sin introducir el segundo postulado de Einstein! No cabe duda de que **no es** una mera cuestión de gustos deducir las transformaciones de Lorentz sin el segundo postulado de Einstein. La veracidad de este postulado se puede comprobar siempre a menos de un error experimental. La disminución de este error podría tener como consecuencia, en principio, la invalidación del postulado en un experimento futuro. El enfoque del joven eludía esto, presuponía menos y permitía deducir la existencia de una velocidad límite. Enviamos entonces el trabajo a una revista norteamericana primero, y a una revista europea después. Ambas lo rechazaron pero con una decisión dividida (siempre algún árbitro era favorable y otro no. El desempate justificaba el rechazo). Uno de los árbitros de la revista europea envió un comentario exhaustivo sobre antecedentes que no conocíamos. Los obtuvimos y resultaron muy valiosos para poner todo el asunto en perspectiva. Pero ni J. D. J. ni los seis árbitros que (además de él) valoraron el trabajo, parecen haberse dado cuenta de lo que finalmente resultó ser el punto débil del trabajo: una suposición sobre la acotación de cierta función, que no era obvia y requería una justificación.

[311] J. Ortega y Gasset "*La rebelión de las masas*", capítulo 12: La barbarie del "especialismo", Altaya, Barcelona, 1993.

científicos empujan el progreso general de la ciencia encerrados en la celdilla de su laboratorio, como la abeja en la de su panal…"

La especialización creciente se ha acompañado de numerosos esfuerzos de divulgación de los diversos resultados especializados obtenidos en cada una de las ciencias y en la técnica. En particular, se dispone de varios trabajos de síntesis divulgativa realmente notable, ya sea en forma de libros[312], ya sea en forma audiovisual a través de programas de televisión o vídeos.
Luego de la Segunda Guerra mundial emergió la teoría de sistemas y se desarrolló con ímpetu durante la segunda mitad del siglo pasado, con vocación unificadora, al igual que la filosofía de sistemas[313].
Pese a todos estos esfuerzos para contrabalancear los efectos de una especialización excesiva, para el común de los científicos y los técnicos se hace cada vez más difícil compartir una concepción del mundo compatible con los conocimientos actuales.
Ni siquiera hoy en día hemos podido controlar las consecuencias de esta dinámica derivada de la especialización, pese a haber dispuesto de un diagnóstico certero y temprano, y pese a que se han hecho esfuerzos (que parecían bien encaminados y fueron parcialmente exitosos) para contrarrestar sus efectos negativos. La organización socioeconómica actual y los valores asociados no son ajenos a ese fracaso.

Es contra este telón de fondo que consideraremos algunos aspectos psicológicos de la genialidad, que sin duda poseían tanto los fundadores de la física relativista como los fundadores de la física cuántica.
La genialidad se puede vincular con cierto grado de desmesura, desmesura que por lo general es ajena al tipo de científico al que alude Ortega en su libro sobre la rebelión de las masas.
Este tipo de científico ciertamente posee una preparación rigurosa, inteligencia y talento, disposición para investigar, e incluso pasión por ello, pero en el fondo es en mayor o menor medida una persona común.
Sus aportes, como los de las abejas obreras, trabajando en una colaboración ordenada, contribuyen a añadir celdilla tras celdilla al panal del conocimiento científico durante sus etapas de desarrollo apolíneo. Unos pocos, dotados de esta naturaleza equilibrada, efectúan aportes en extremo importantes en el marco de un paradigma establecido y sobresalen. Un puñado se transforma en personajes mediáticos.

Pero los cambios de paradigma se deben por lo general a personalidades dionisíacas dominadas por violentas pasiones e intuiciones repentinas, personas cuyos rasgos se parecen más a la de los poetas o los profetas que a los científicos apolíneos.
Anhelo puro de saber junto con ambición y afán de poder, susceptibles de organizarse como grandes pasiones y de conducir a conductas que, vistas desde un punto de vista ético, pueden aparecer como reprobables.

---

[312] Por ejemplo, el de P. Cloud "*El cosmos, la Tierra y el hombre: breve historia del Universo*", Alianza, Madrid, 1981.

[313] Algunas de las ideas más importantes de la teoría de sistemas se encuentran en el libro de Ludwig von Bertalanffy "*Teoría General de los Sistemas: Fundamentos, desarrollos y aplicaciones*", Fondo de Cultura Económica, Méjico, 1976. La importancia de la idea de sistema para la filosofía es remarcada por Mario Bunge en todos sus libros. Pueden consultarse: "*Teatrise on Basic Philosophy*", volumen 4: A World of Systems, Reidel, Boston, 1979 y "*Emergencia y Convergencia*", Gedisa, Barcelona, 2004.

Personas, por qué no decirlo, que presentan algunos rasgos que comúnmente se asocian con la locura. Inclusive, algunas veces manifiestan en algún momento de su vida una conducta que conduce a que sean tratados como enfermos mentales.

Las ideas sobrevaloradas o preconcebidas por el científico genial son parte de su fuerza, al igual que una intensa actividad combinatoria de pensamiento que vincula e interpreta buena parte de los incidentes de la vida como apoyos a una hipótesis dominante, que tiende a excluir de la conciencia los contenidos que para ese propósito son irrelevantes[314].
El estrechamiento del campo mental que acompaña la sobrevaloración de una idea se manifiesta en una formidable concentración de la atención y tiene como consecuencia una proverbial distracción, cuyas consecuencias son rayanas en lo ridículo.
Todo esto no es posible sin pasiones extraordinarias, sin desmesura, sin orillar la locura o precipitarse en ella. Aunque la mayor parte de los documentos que pueden arrojar alguna luz sobre las características psicológicas de los investigadores geniales tiende a ocultar estos aspectos que la gente común tiende a considerar infamantes, la evidencia disponible sugiere que la descripción anterior se ajusta bastante bien a los hechos[315].

Hay un punto en particular que conviene plantear para finalizar esta parte. Además constituye una conexión natural con la siguiente: se trata de la interferencia de ideales. A medida que la humanidad evoluciona va añadiendo nuevos ideales, conciliables solo en parte. Como consecuencia se tiene una moral conflictual[316], particularmente evidente en las sociedades que han recibido una mayor influencia de la tradición judeocristiana.
Cuando esto se tiene debidamente en cuenta, permite comprender los conflictos éticos que acompañan toda labor científica socialmente destacada y en particular la que se ha consagrado ya como genial.
No parece razonable ni conveniente escandalizarse por ciertas limitaciones o defectos: vivir conduce a situaciones en las que se tienen que sacrificar parte o todos los ideales que se poseen y se sienten. Sobresalir lo exige aún más. Este pudo haber sido el caso durante la imposición del nuevo paradigma asociado con la teoría de la relatividad. También podría ser, a la inversa, el caso actual, en el que al parecer se protegen y mantienen a ultranza ciertos paradigmas en las ciencias biomédicas, gracias a un sistema de arbitraje de las propuestas de investigación en el que muchas veces la decisión pasa por personas cuyos intereses económicos y cuya notoriedad está asociada a emprendimientos basados en los paradigmas amenazados[317].
Como decía Carlos Vaz Ferreira, desde su óptica característica:

---

[314] Una vez le preguntaron a Rutherford como había logrado encontrar las leyes del decaimiento radioactivo y había podido determinar la composición de las radiaciones nucleares. Respondió: "No pensando en otra cosa durante siete años."

[315] Ver, por ejemplo, el libro de E. Kretschmer "Hombres geniales", capítulo 8, Labor, Barcelona, 1954.

[316] Sobre la incompatibilidad entre ciertos valores se pueden consultar los libros de I. Berlin, "Contra la corriente", Fondo de Cultura Económica, México, 2006 y "Conceptos y categorías", Fondo de Cultura Económica, Madrid, 1983.

[317] G. Moran "*Silencing scientists and scholars in other fields: power, paradigm control, peer review, and scholarly communication*", Ablex Publishing Corporation, Greenwich, CT, 1998.
B. Martin "Suppression of dissent in science", Research in Social Problems and Public Policy, vol.7, 1999, pp.105-135.

“Se podría concebir un hombre que tuviera tanta caridad como los santos de la historia, tanto patriotismo como sus héroes, tanto amor a la ciencia como los mártires de la verdad; que tuviera todos los sentimientos en su máximo histórico, y, además, en su máximo también los sentimientos no históricos: sentimientos de familia, de amistad hacia los otros. Difícilmente podría su actitud ser histórica. Desde luego, a la historia va lo que ciertos grandes hombres hicieron; no puede ir lo que otros, quizás más grandes todavía, se inhibieron de hacer.” [318]

15.5 La sociedad post-moderna y la relatividad.

El agregado y la interferencia de nuevos ideales en sociedades que a partir de fines de la década de 1980 han sufrido un segundo y más completo proceso de globalización [319] han aumentado el desconcierto y la confusión. No obstante, no han impedido que algunas personas intentaran esbozos de una moralidad mundial, por cierto muy tenue, pero que en principio parecería pasible de ser implementada[320].
Las raíces de las dificultades que es necesario afrontar cuando se intenta acordar al respecto de una ética mundial no parecen encontrarse propiamente en la época actual, pese a que se presenta como meramente reactiva, anti-utópica, con escasos ideales fuera del ámbito de lo individual y personal, con escasa fe en la razón y en la ciencia, aunque adore la tecnología.
Pero ocurre que, aún como integrantes de la sociedad globalizada contemporánea no podemos permanecer indefinidamente sin saber quiénes somos.
Por este motivo la principal inquietud compartida de nuestra época se relaciona con la búsqueda de algún modelo que responda a esta interrogante.
La clave de las dificultades actuales parece encontrarse más bien en la etapa final de la época precedente, la época moderna, y en una visión del mundo que se constituyó en torno a la ciencia.
Desde la antigüedad clásica, pasando por la edad media, hasta el inicio de la edad moderna en Europa, se pensaba en el Universo como en una gran cadena del ser.
Una estructura jerárquica con el plano humano en un nivel intermedio, en la que se puede ascender progresivamente desde los planos inferiores hasta la cúspide, el *ens perfectissimum*, en el que lo verdadero, lo bueno y lo bello alcanzan su plenitud.
Una simplificación popular de esta organización resumía la cadena en la progresión hacia los cielos, partiendo de los infiernos y pasando por el plano terrestre.
Pero, en todo caso, lo que estaba dotado de más realidad no era accesible a través de los sentidos. No se encontraba en nuestro plano ontológico, ni en el plano situado por debajo de él, sino por arriba de él. Se descubría mediante símbolos, que situados siempre en un plano inferior, apuntaban a realidades ubicadas en planos superiores.

Lo que se suele aludir como la concepción moderna del mundo comenzó a tomar forma durante el siglo XVII y se disolvió durante la primera mitad del siglo XX.

---

[318] C. Vaz Ferreira “¿Cuál es el signo moral de la inquietud humana?” en “*Estudios filosóficos*”, Aguilar, Buenos Aires, 1961.

[319] Un primer proceso de globalización, más débil, ocurrió durante los veinte años previos a la Primera Guerra Mundial, asociado con el predominio inglés y en general, europeo.

[320] H. Küng, “*Proyecto de una ética mundial*”, Planeta-de Agostini, Barcelona, 1994.

Se había organizado en torno a una visión desde las ciencias y evolucionó hasta hacer del conocimiento científico el ideal de todo conocimiento.
Se distingue un micro-cosmos (de las células biológicas a los electrones, pasando por los átomos y sus núcleos), un meso-cosmos a escala humana (con cuerpos materiales cuyas dimensiones van de los milímetros a los kilómetros), y un macro-cosmos que comprende en su interior a todo lo demás (incluido nuestro planeta y el resto del sistema solar, los otros soles, lunas y planetas de nuestra galaxia, las demás galaxias de nuestro cúmulo, todos los cúmulos, los super-cúmulos de galaxias y las enormes burbujas de vacío entre ellos).

Las ciencias reconocen un único plano ontológico, material y cuantificable, cerrado en sí mismo.
Se tendió a reducir lo real a lo material o al menos a considerar todo lo demás como meros epifenómenos.
Dentro del ámbito de lo material, se intentó reducir los procesos y estructuras del macro y el meso-cosmos a los procesos y estructuras en el micro-cosmos.
La gran cadena del ser fue eliminada, junto con el tipo de actitud contemplativa que conllevaba, a favor de un intento de dominio sobre la naturaleza.

Una actitud faústica, una vocación técnica, comenzó a dominar las actividades de los pueblos de Europa y a incidir sobre las manifestaciones de la cultura en el continente y sus colonias.
Un control eficaz de la naturaleza requiere una elevada capacidad de predicción de los acontecimientos.
Una elevada capacidad de predicción requiere a su vez la capacidad de descubrir leyes apropiadas para ese propósito.
El descubrimiento de ese tipo de leyes solo fue posible en un marco específico de conocimiento objetivo, obtenido aplicando el método experimental-hipotético-deductivo-experimental basado en última instancia en el número y la medición.
Junto con el progresivo predominio de la ciencia y de la técnica basada en la ciencia, se reconfiguraron los valores en una cierta jerarquía que permitió dotar de sentido a la idea, típica de la modernidad, de un progreso sin fin, no solo en lo referente al conocimiento objetivo, sino en todos los aspectos de la vida humana que no se habían desvanecido con la época anterior.
La concepción moderna acompañó el crecimiento, la diversificación y el extrañamiento recíproco de las ciencias, hasta que, finalmente, cuando la ciencia del siglo XX tomó distancia y disolvió lo que parecía ser el sustrato firme en el que se basaban las ideas sobre el mundo físico, la concepción moderna del mundo perdió vigencia.

La modernidad dio paso a la post-modernidad actual.
En buena medida esta última es una respuesta al desencanto causado por el fracaso en el cumplimiento de las promesas de progreso en lo social, en lo político, en lo económico e inclusive en el conocimiento científico mismo que se podía constatar al final de la primera mitad del siglo XX.
La gente de la época moderna era optimista porque suponía que sabía más que sus predecesores y porque tenía fe en un progreso ilimitado.
La gente postmoderna ya no lo es tanto: supone que sabe más porque ha descubierto lo poco que puede llegar a saber.

La época postmoderna es, pues, una época de desencanto, de renuncia a las utopías y a la idea de progreso de conjunto típico de la época moderna: pone énfasis en la competencia por el desarrollo individual más que en la cooperación solidaria.
Se comienza a pensar que la aventura de la ciencia no es aquella empresa sin límites destinada a la generación de conocimiento verdadero, acumulativo y de validez universal, en cuya fe se sustentó la época anterior.
Por lo general los post-modernos no tienen fe en otro progreso ilimitado que no sea, propiamente, el de la técnica y las comodidades que trae aparejadas para la vida diaria (aunque en muchos casos no puedan acceder a ellas).
Y eso cuando se olvidan de lo que puede deparar un futuro incierto y un mundo que perciben lleno de conspiraciones.

Como consecuencia, a diferencia de las épocas pasadas, no parece haber una visión integrada ni un sentido único de la realidad que caractericen a nuestra época.
En relación con este último asunto, resulta interesante la siguiente observación de Huston Smith[321]:
"Heidegger creía que las eras son impulsadas por obras de arte que reúnen prácticas diseminadas, convirtiéndolas en modelos unificados y persuasivos de comportamiento, mostrándoselos al público para que puedan relacionarse entre sí según las pautas que ejemplifican. No obstante, pensó que nuestra época era la primera cuyo paradigma no era una obra de arte, sino de tecnología: la central hidroeléctrica. Al ser una construcción tecnológica, carece de valores morales. La central eléctrica convierte la potencia del río en algo disponible para cualquier propósito. Como la electricidad puede utilizarse para satisfacer cualquier deseo que podamos albergar, este paradigma no proporciona ninguna directiva ni motivación de cara a la acción. La consecuencia es un vacío de significado y propósito…"

Estas notas características de la postmodernidad se manifiestan en los trabajos científicos contemporáneos a través de:

-La ruptura del paradigma de progreso acumulativo y lineal, que en la época moderna se consideraba propio del conocimiento científico. Esto se manifiesta en la reaparición de temas que se consideraban definitivamente cosa del pasado[322].

-El descenso en la calidad promedio en la producción científica que se envía para su publicación[323]. Otros fenómenos, en parte pero no exclusivamente relacionados con el

---

[321] H. Smith "*Más allá de la mente postmoderna*", Parte I: el bosque oscuro, Kairós, Barcelona, 2001.

[322] En física se reabrieron los debates epistemológicos en torno a la mecánica cuántica, volviendo a discutir aspectos que la mayor parte de los físicos consideraban ya resueltos a fines de la década de 1930. Algunos de estos aspectos pueden verse en los libros de K. Popper "*Teoría cuántica y el cisma en física*", Tecnos, Madrid, 1985, de C. von Weizsäcker, Th. Görnitz y H. Lyre "*The structure of Physics*", Springer, Dodrecht, 2006, y de M. Guerra "*El enigma cuántico*", Editorial Grupo Magró, Montevideo, 2009. En biología, sobre todo desde el inicio del nuevo siglo, se comienza a notar un interés renovado por las propiedades coloidales del citoplasma y sus modificaciones estructurales (fenómenos cooperativos manifestados como cambios de fase), en parte retomando algunos conceptos que habían desaparecido de los textos y continuando trabajos del primer tercio del siglo pasado. Sobre esto último se puede ver, por ejemplo, el libro de G. Pollack "*Cells, gels and the engines of life: a new, unifying approach to cell function*", Ebner & Sons, Seattle, USA, 2001.

descenso en la calidad, son el incremento en los redescubrimientos y en las reiteraciones, que más a menudo de lo que se cree, pasan desapercibidos.

-La pérdida de pertinencia de los resultados de la investigación científica, no solamente para el público en general, sino para los científicos mismos.[324]

Podría suceder, como teme J. Lévy-Leblond[325], que en un futuro no demasiado lejano, apoyada por una opinión pública desencantada con los aspectos más especulativos de las ciencias básicas, termine por imponerse una suerte de tecno-ciencia cuya creciente eficacia práctica opaque progresivamente el quehacer científico tradicional.

Pero detenernos en este futuro posible nos aleja del tema que aquí nos ocupa. Regresemos entonces a la teoría de la relatividad.
Esta teoría pertenece a una etapa de transición entre dos épocas: se puede considerar ya sea como la última gran teoría física de la época moderna (en ese espíritu trabajaron sus fundadores), ya sea como un primer síntoma de lo que se avecinaba. La mecánica cuántica, por su parte, marca ya la entrada de la post-modernidad a través de la ciencia.

Consideremos brevemente el efecto de la teoría de la relatividad en la disolución de las ideas sobre el mundo físico propias de la época moderna.
La teoría restringida introduce el concepto de continuo tiempo-espacio como un absoluto, y al hacerlo relega a los conceptos de espacio tridimensional y de tiempo unidimensional a meros conceptos relativos al observador (marco inercial) y que se modifican al pasar de un observador a otro en movimiento rectilíneo y uniforme respecto del primero.
Con esta relativización del tiempo y del espacio se contradicen las nociones intuitivas en base a las cuales percibimos el mundo como un espacio visual tridimensional independiente de una dimensión temporal, también presente como un flujo absoluto de eventos ordenados por un antes y un después.

---

[323] Esto es una percepción común a los árbitros con muchos años de experiencia en valorar trabajos de investigación presentados para su publicación en revistas científicas y en libros de procedimientos de congresos, aún en los que se sabe de antemano que mantienen estándares elevados de exigencia. No se escapan a esta tendencia los cada vez más frecuentes trabajos provenientes de equipos numerosos de investigadores, trabajos que cabría esperar fueran revisados por cada uno de los coautores.

[324] Por supuesto, cuando se trata de temas ajenos al tema específico, y por lo general estrecho, que cada científico se encuentra investigando. De acuerdo con estudios del Institute of Scientific Information (www.isi.edu), dos tercios de los artículos científicos no son citados jamás por otros investigadores. A su vez, las citas de artículos en trabajos de investigación publicados se remontan por lo general hasta cinco años hacia el pasado, raramente diez o más. La postmodernidad en las ciencias se caracteriza por un crecimiento exponencial en el número de publicaciones y al mismo tiempo por una creciente pérdida de visibilidad y rápida obsolescencia de esas publicaciones. Esto ocurre pese a los esfuerzos para revertir la situación, creando bases de datos comprehensivas y disponibles en internet, e implementando numerosas redes de investigadores.

[325] Sobre la incidencia de la postmodernidad en la ciencia se puede consultar la conferencia de J. Lévy-Leblond "(Re) mettre la science en culture: de la crise épistémologique à l'exigence éthique", presentada en Palermo, en marzo de 2007, durante la inauguración del Institute for Scientific Methodology; www.i-sem.net.

En particular, la teoría restringida predice que el orden temporal de dos eventos puede invertirse al pasar de un marco de referencia inercial a otro, siempre que no tengan una conexión causal entre ellos (es decir, siempre que no se pueda hallar una línea de universo en el continuo tetra-dimensional formado por el tiempo-espacio que los incluya a ambos).
La confusión creada en relación con los nuevos conceptos de tiempo y espacio se manifestó con el planteo de numerosas paradojas. La más famosa de estas paradojas, la de los gemelos, encontró su lugar en la literatura y el cine, y ha alimentado la fantasía de varias generaciones. [326]

La relatividad del espacio y del tiempo fue interpretada por parte del público como un apoyo a las posturas filosóficas típicas del relativismo, proveniente de la ciencia.
Einstein mismo puede haber contribuido involuntariamente a esta confusión. Cuando los periodistas le preguntaban a qué se debía su extraordinaria popularidad, solía contestar que se debía a la relatividad y a la manera en que se la interpretaba: tal o cual cosa es relativa, todo es relativo.
Así, mal comprendida, la teoría de la relatividad puede haber contribuido en la formación del concepto post-moderno de verdad como una cuestión de perspectiva o de contexto más que como algo universal.

Pero pese a todo, las consecuencias de cambiar el marco de referencia que predecía la teoría restringida podía comprenderse estableciendo una analogía con un cambio de perspectiva visual para contemplar el mismo objeto.
En la década de 1890, algunos investigadores fotografiaban el mismo objeto en movimiento, simultáneamente con varias cámaras a varias distancias y desde varias direcciones. Vitz y Glimcher pensaban que, si cada cámara se asimila a un observador, la estructura de la representación resultante evoca un análisis relativista del espacio y el tiempo[327].
El movimiento cubista en pintura, desde su aparición en 1907, intentó presentar en el lienzo la misma realidad vista desde múltiples puntos de visión. Posteriormente, cuando la teoría restringida de la relatividad se hizo más conocida y ya había aparecido la teoría generalizada, algunos de los pintores cubistas con inclinaciones filosóficas, como Salvador Dalí, intentaron vincular el cubismo con la teoría de la relatividad.
Parece más verosímil que el cubismo recibiera la influencia de las reflexiones que publicó Poincaré en 1902 (en uno de sus libros ya mencionados, en principio destinado al gran público)[328], sobre la posibilidad de añadir a las tres dimensiones del espacio visual al menos una dimensión espacial adicional basada en la combinación de las sensaciones táctiles y cenestésicas asociadas a la exploración de los cuerpos materiales.
Las opiniones de Guillaume Apollinaire [329] sobre la cuarta dimensión en las artes plásticas refuerzan esta presunción.

---

[326] Un planteo adecuado de esta paradoja requiere el empleo de marcos de referencia acelerados, de modo que el marco natural para estudiarla no es la relatividad restringida sino la generalizada. En este trabajo la hemos ignorado ex profeso.

[327] P. Vitz y A. Glimcher "*Modern art and modern science*", Traeger, New York, 1984.

[328] H. Poincaré "*La Science et l'hypothèse"*, Flammarion, París, 1902.

[329] Guillaume Apollinaire "*Les peintres cubistes*", Figuière, Paris, 1913.

La teoría generalizada de la relatividad resultó aún más disolvente de los preconceptos propios del sentido común. Ahora la analogía con el cambio de perspectiva visual solo se puede aplicar localmente. Hacerse una idea intuitiva sobre el tiempo-espacio-materia curvado como un todo es algo que parece estar fuera de las posibilidades de una persona. Solo haciendo un análisis abstracto como el presentado en los capítulos 9 a 11, se puede obtener una descripción adecuada, que escapa a las representaciones gráficas y a todo lo que nos es familiar en nuestra experiencia cotidiana.
Cuando este análisis no resulta accesible, parecería que lo único que queda por hacer es conformarse:
"La reacción normal ante una primera toma de contacto con la relatividad es: <<Me parece que lo comprendo, lo único que ocurre es que no lo puedo creer>>.
Se trata de una teoría con la que los físicos suelen tardar unos cinco años en familiarizarse, pero no solamente porque sea compleja u oscura, sino simplemente porque es terriblemente extraña.
La regla del juego consiste en aceptar las consecuencias de los postulados sin preocuparse por lo insólito de sus implicaciones y sin preguntarnos cómo pueden ser así.
Lo único que hay que hacer es aceptarlas y ver el modo en que contradice nuestra experiencia.
Así pues, el lector sólo debe tener fe, en la esperanza de que finalmente se evidenciará su coherencia." [330]

El modelo de Universo esférico que Einstein propuso en 1917, ilimitado pero con un radio de curvatura, un volumen y una masa de materia finitos, cuando se divulgó causó cierta perplejidad en el público. El común de las personas daba por sentada la relación entre la materia y el espacio euclidiano propia de la física newtoniana: un espacio ilimitado con una densidad promedio de materia no nula debe contener una cantidad infinita de materia.
No obstante, la teoría generalizada de la relatividad abrió la puerta al desarrollo de la cosmología científica que revisamos en los capítulos 12 y 14. Junto con la mecánica cuántica, impulsó ingeniosas especulaciones[331] sobre universos paralelos, agujeros de gusano, puertas a otros mundos y paradojas apasionantes que, a través de una concienzuda actividad de divulgación, exhiben un notable éxito mediático.

Por otra parte, el modelo plano Λ-CDM ha permitido construir una historia bastante detallada del cosmos en expansión.
Se basa en una narración sobre la evolución de la materia y la energía oscura, desde un hipotético inicio (que como vimos en 12.4.2 se establece con notable precisión) hasta la actualidad.
La narración se apoya en una extrapolación de las leyes físicas conocidas, en suposiciones que se consideran razonables y en asumir en forma explícita la intervención cada vez más frecuente del azar a medida que la estructura en la materia aumenta su complejidad hasta llegar a la aparición de vida en el cosmos.

---

[330] R. March "Física para poetas", Siglo XXI, México, 1995.

[331] K. Thorne "*Agujeros negros y tiempo curvo*", Crítica, Barcelona, 1995.

Se desarrolla encerrada en el plano ontológico de la materia, como todo lo que emana de la ciencia, y constituye la más formidable extensión de la visión horizontal en ese plano alcanzada y admitida hasta el momento.
El acuerdo que existe al respecto es sorprendente, en vista de las violentas controversias que se produjeron en un pasado no tan lejano entre los científicos y filósofos del antiguo bloque soviético y los científicos y filósofos del por aquel entonces denominado bloque occidental, y que ahora ya nadie parece querer recordar[332].

[332] Entre 1983 y 1984 un colega publicó una serie de trabajos sobre una teoría relativista de la gravitación con algunas consecuencias cosmológicas, entre ellas una explicación de los hechos que se atribuyen a la expansión del Universo sin necesidad de esa expansión y en forma compatible con la idea de una materia eterna. Recibió numerosas cartas reprobatorias que trasuntaban fuertes emociones en los autores. Los artículos de mi colega habían sido publicados en revistas científicas internacionales, arbitradas y prestigiosas, como Il Nuovo Cimento y los Annalen der Physik (la misma revista en la que Einstein publicó su famoso trabajo de 1905 sobre la relatividad especial) de modo que tuvieron amplia difusión. Si hubiera podido publicar eso mismo veinte años antes, hubiera recibido reprobaciones es cierto, pero también apoyos decididos y entusiastas provenientes de uno de los bloques.

16. Epílogo: El principio de relatividad y el problema del conocimiento

"Dondequiera que uno esté, piensa que ése es el centro; el centro del mundo está en todas partes, y por tanto en ninguna, y su circunferencia no está en ninguna parte"
Nicolás de Cusa (1401-1464), en Docta Ignorancia.

"Las ciencias naturales clásicas consolidaron la impresión de que la mente humana adulta es un absoluto, capaz de descubrir leyes físicas absolutas que tratan de una materia absoluta contenida en un espacio y un tiempo absoluto. La escisión cartesiana entre mente y materia es la piedra angular donde se apoya esta concepción."
M. Delbrück "Mente y materia: ensayo de epistemología evolutiva", Alianza Editorial, Madrid, 1986.

"*Cerebros complejos pero muy diferentes del nuestro, productos de historias naturales muy diferentes de la nuestra, pueden llegar a concebir ideas acerca de la "realidad" que no tendrían sentido para nosotros, pero que en el contexto foráneo en que se desarrollan son tan útiles para esos seres como las nuestras lo son para nosotros."*
G. Munévar en A. Rivadulla (Ed.) "Hipótesis y verdad en ciencia: ensayos sobre la filosofía de Karl Popper", Editorial Complutense, Madrid, 2004.

"Es indudable que los entresijos de la física necesitan un buen limpión…Empieza a no verse clara la relación entre la doctrina a que se ha llegado y la realidad. De un lado están los grandes hechos observados, de otro el aparato hiper-tenue de las teorías, telas de araña sutilísimas, como espectrales, reducidas casi a puras abstracciones de simbolismo matemático. La correspondencia entre estas teorías y aquellos hechos, entre el cuerpo de observaciones y el cuerpo de los conceptos o doctrinas se ha hecho equívoca. Hay, sin duda, correspondencia, pero no se sabe muy bien en qué consiste.
A veces parece como si lo que la teoría dice no tuviera nada que ver con las cosas…Ciencia parece querer decir conocimiento, y conocimiento parece significar presencia en nuestra mente de lo que las cosas son. Más la ciencia física no nos pone en la cabeza más que fichas…De las cosas mismas no pasa nada o pasa poco más que nada a nuestra mente. ¿Cabe seguir llamando a eso conocimiento? ¿No podría, con igual fundamento, llamársele guardarropía?"
J. Ortega y Gasset, "Bronca en la física", Obras completas, volumen 5, Alianza, Madrid, 1987.

"Pero al igual que ocurre con la mayoría de las cosas, la ciencia tiene dos caras. Si creemos que lo que nos muestran sus instrumentos agota toda la realidad posible, acabaremos abocados al cientificismo y al materialismo. Si, por otro lado, aceptamos que su visión es en principio limitada, podemos dirigirnos a ella no tanto en busca de pruebas como de indicios acerca de la naturaleza de lo que queda fuera de su alcance"
H. Smith "La verdad olvidada", Kairós, Barcelona, 2000.

"Despreciar la filosofía cuando se está involucrado en una reformulación de ideas es admitir la absoluta corrección de los prejuicios filosóficos que puedan haber introducido una nana o un maestro de escuela o propios de los modos de expresión corrientes."
Alfred N. Whitehead "The principle of relativity with applications to physical science", Cambridge University press, U.K, 1922.

La importancia de los cambios producidos en la formulación y en la comprensión del alcance del principio de relatividad se resume muy bien en la cita de Nicolás de Cusa con la que se inicia el presente capítulo.
Vista desde una perspectiva más amplia, la historia del principio de relatividad del movimiento es solo un aspecto de la historia de la progresiva toma de conciencia de que tanto en el mundo físico como en el sistema formado por las culturas humanas, no hay propiamente un centro. Cuando salimos a la caza de lo real, lo hacemos siempre desde una circunstancia y con una perspectiva que varían de una persona a otra, de una época a otra y de una cultura a otra. Nuestra aldea es eso, una aldea, y como tal debemos aceptarla, vivir en ella e intentar, hasta donde nos sea posible, eliminar en lo percibido y lo pensado, aquello que podría ser meramente accidental y subjetivo.

Pero en el estudio del mundo físico relatividad no es relativismo: no todo es relativo. Hemos visto que se pueden identificar invariantes sobre los que, en principio, todos los observadores pueden estar de acuerdo: por ejemplo, los eventos físicos, tales como las desintegraciones de núcleos atómicos y algunas magnitudes asociadas a los eventos, como el intervalo que los separa en el tiempo-espacio.
La posibilidad de identificar invariantes permitió formular algunas de las leyes más importantes de las ciencias de la naturaleza. Tanto la teoría de la relatividad, como la física que vino después (y que fue posible gracias a aquella), continuaron enfatizando las leyes de conservación y las características invariantes que se pueden atribuir al mundo material, como lo había hecho ya la física previa.

La búsqueda de absolutos es una necesidad psicológica. Imprime una dirección al quehacer de los científicos, los artistas y los filósofos.
Adoptando una postura filosófica realista, es posible conjeturar que esta exigencia de nuestra condición humana tiene un correlato en la realidad, en una unidad más profunda que subyace a la diversidad de puntos de vista y a la multiplicidad de los fenómenos.
Pero para el trabajo de investigación en cada ciencia particular, no es indispensable tener en cuenta, ni siquiera entrever, ese aspecto conjetural de la realidad.
De hecho ni siquiera es necesario preocuparse por el conocimiento de la realidad.
Por ejemplo, el punto de vista del instrumentalismo consiste en poner el foco en el mayor o menor éxito de un modelo teórico para calcular y predecir los resultados de los experimentos, sin asumir que por ello un modelo sea más o menos verdadero que cualquier otro. Esta última cuestión la descarta como carente de sentido.
De una u otra forma, el instrumentalismo ha sido suficiente como marco conceptual para una parte del trabajo de muchos intelectuales desde hace 2000 años.
Fue adoptado por físicos eminentes como Dirac, durante el siglo XX. Permite eludir o minimizar la importancia de preguntas tales como la que preocupaba al no menos eminente Wigner: ¿cuál es el motivo profundo de que la matemática resulte tan poderosa como herramienta para las investigaciones en las ciencias naturales?

16.1 Instrumentalismo, realismo y relatividad del movimiento.

El Cardenal Bellarmino adoptó una posición que ahora tildaríamos de instrumentalista, durante su primer enfrentamiento con Galileo en relación con los movimientos de los cuerpos celestes.
Galileo, como posteriormente Newton, tenían un concepto absoluto del movimiento. Ambos adoptaron, como Einstein, el realismo como posición filosófica.
Galileo pensaba que la Tierra se movía **realmente** alrededor del Sol, el cual estaba

inmóvil en el cosmos.
Bellarmino, en 1615, argumentó en contra sobre la base de los trabajos de 1450 de Nicolás de Cusa sobre el movimiento. Este último supuso que si el espacio era infinito, entonces no se puede introducir un criterio de reposo o de movimiento basado solamente en el espacio: el reposo o el movimiento tenía un sentido relativo, el movimiento se podía determinar solamente respecto de los cuerpos materiales.
Como consecuencia se podía describir el movimiento de los cuerpos celestes tanto respecto de la Tierra como respecto del Sol.
Por el contrario, cuando se admitía que el cielo era una esfera finita, era posible describir respecto de este cuerpo un movimiento verdadero.

Dice von Weizsäcker [333], siendo Bellarmino capaz de pensar en la relatividad del movimiento, quería que Galileo presentara el sistema copernicano como una hipótesis (un modelo, en sentido moderno) y no como una verdad.
Galileo lo admitió sin convencimiento, hasta que 15 años después, no pudo ocultarlo más y se produjo el conocido desenlace de arresto domiciliario y prohibiciones.
Ahora, después de disponer de la teoría de la relatividad y de la cosmología física relativista, podemos darnos cuenta de en qué sentido se puede afirmar que Galileo llevaba la razón, desde una postura ontológica realista.
Según vimos en 12.4 y 14.5.2, el flujo de Hubble en el tiempo-espacio permite definir un marco inercial local en co-movimiento con la materia.
En este complejo y sofisticado marco de conceptos teóricos es razonable definir (y puede medirse) el movimiento de nuestra galaxia respecto de ese marco inercial local, así como el movimiento del Sol respecto de la galaxia y el movimiento de la Tierra respecto del Sol.

Pero en la época de Galileo primero y de Kepler después, los argumentos se basaban en:
(1) La mayor simplicidad de la descripción heliocéntrica.
(2) Que parecía más razonable que el cuerpo mayor (eso se sabía desde Aristarco) y más luminoso fuera el centro en lugar de serlo la Tierra.
(3) La sospechosa coincidencia del promedio a largo plazo de los períodos de rotación de Mercurio y de Venus alrededor de la Tierra con la duración del día solar en el sistema geocéntrico comparada con su explicación en base a encontrarse más cerca y girar más rápido que la Tierra respecto del Sol en el sistema heliocéntrico.
(4) La mayor simplicidad de las elipses que Kepler postulaba para las órbitas planetarias en el sistema heliocéntrico respecto de las curvas atrabiliarias que esos mismos planetas describían en la bóveda celeste en el sistema geocéntrico.

Finalmente, con la dinámica de Newton y su ley de gravitación universal, el análisis del movimiento y sus causas en base al sistema heliocéntrico mostró su superioridad respecto del viejo sistema geocéntrico.
No obstante, desde el punto de vista del instrumentalismo, la pregunta sobre cuál de los dos modelos, el geocéntrico o el heliocéntrico, es el verdadero carece de sentido, y desde ese punto de vista Bellarmino llevaba la razón. [334]

---

[333] C. von Weizsäcker, Th. Görnitz y H. Lyre "*The structure of Physics*", parte 2.8, Springer, Dodrecht, Holanda, 2006.

[334] A Bellarmino no le preocupaba si el movimiento era o no absoluto. Como creyente y como teólogo le importaba la revelación. Ver la cita que aparece en el libro de G. de Santayana "*The crime of Galileo*",

Como señala Mittelstaedt [335] en el debate entre los partidarios de la concepción relativa (Leibniz, Mach, Einstein) y los partidarios de la concepción absoluta del movimiento (Galileo, Newton) hay un trasfondo filosófico en el que emergen dos concepciones diferentes de la relación entre la materia y el espacio.
Si sobre la base de una concepción dualista, se supone que hay un espacio y una materia, considerados como si fueran dos realidades independientes, se puede hablar de movimiento absoluto de los cuerpos materiales respecto de ese espacio.
Pero si se considera que la extensión es un predicado esencial de la materia, considerada como sujeto, que la extensión es una característica fundamental de la materia, entonces las relaciones espaciales entre los cuerpos son una consecuencia de su extensión.
La geometría física no se puede desvincular de la materia: el espacio se subordina a la materia y el movimiento absoluto no tiene sentido en esta concepción monista.
La manera en que se dio una solución a este problema del movimiento en la teoría restringida, en la teoría generalizada de la relatividad y en la cosmología relativista ya fue estudiada en los capítulos precedentes.

16.2 Relatividad de las categorías del conocimiento [336].

No obstante, el dualismo encerrado en las lenguas indoeuropeas (sujeto y predicado, sustancia y atributo) no se puede hacer desaparecer del pensamiento científico por más monista que pueda ser la intención del investigador.
Además, estos lenguajes enfatizan el tiempo (a través de enunciados que distinguen claramente entre pasado, presente y futuro). El concepto de tiempo ocupa un rol central en el desarrollo del pensamiento occidental, sobre todo en la época moderna.
Los estudios de lingüística comparada muestran un panorama muy diferente. Esto nos conduce a considerar la relatividad de las categorías del conocimiento desde una perspectiva lingüística, investigada y enfatizada hasta la polémica, en el segundo cuarto del siglo XX, por Benjamín Whorf y otros investigadores.

Por ese entonces se completaron los trabajos de von Uexküll y otros biólogos que mostraban y enfatizaban la relatividad biológica de las categorías del conocimiento, entre ellas las formas de organizar las sensaciones visuales y cenestésicas en percepciones espaciales y temporales, o sea al espacio y el tiempo como formas de la intuición. Sobre todo esto las neurociencias tienen mucho que decir, en particular la neurociencia cognitiva.

Mientras que los aspectos básicos de la percepción dependen de un aparato psicofísico

---

p.99, The University of Chicago Press, Chicago, 1955.

[335] P.Mittelstaedt "*Problemas filosóficos de la física moderna*", capítulos 2 y 4, Alhambra, Madrid, 1969, y "Der Dualismus von Feld und Materie in der allgemeinen Reltivitätstheorie" en N. Nelkowski (Ed.) "*Einstein Symposium Berlin*", Springer Verlag, Berlín, 1979.

[336] Ver L. von Bertalanffy "*Teoría General de los Sistemas: Fundamentos, desarrollos y aplicaciones*", capítulo 10, Fondo de Cultura Económica, México, 1976; M. Delbrück "*Mente y materia: ensayo de epistemología evolutiva*", Alianza Editorial, Madrid, 1986; K. Popper "*Conocimiento objetivo: un enfoque evolucionista*", Tecnos, Madrid, 1998; A. Rivadulla (Ed.) "*Hipótesis y verdad en ciencia: ensayos sobre la filosofía de Karl Popper*", Editorial Complutense, Madrid, 2004.

común a nuestra especie, el pensamiento simbólico está fuertemente condicionado por la cultura.
Aún nuestra matemática y nuestra lógica están fuertemente condicionadas por la estructura de nuestro lenguaje y de nuestros mecanismos perceptivos, incluido el predominio de la percepción visual en nuestra especie.
Piénsese tan solo en el desarrollo del cálculo infinitesimal y el análisis matemático.
En un comienzo, una curva dibujada en un papel, junto con el esbozo de unos ejes de coordenadas. Enseguida una modalidad de descripción: visualizar la recta tangente en cada punto y seguir la variación de su pendiente al pasar de un punto a otro.
Después, el cociente de diferencias, la derivada y la diferencial expresados en forma analítica.
Luego las funciones de dos varias variables, el plano tangente en cada punto y su variación de un punto a otro, acompañados de las derivadas parciales y la diferencial en varias variables. Finalmente, unos pasos más adelante en la abstracción hasta las aplicaciones entre espacios de Banach y la diferencial de Frechet.
El escoger la recta como curva de referencia no ha sido un accidente: como mostraron Hubbel y Wiessel [337], es consecuencia de la forma jerárquica en que nuestro cerebro organiza las señales visuales provenientes de la retina. Pero una vez que se emplea la recta en una descripción apoyada en un croquis plano, condiciona todo el proceso de elaboración de abstracciones futuras, hasta el cálculo infinitesimal en espacios de dimensión infinita.

Vistas las cosas desde una perspectiva biológica, se puede presumir que las poblaciones de homínidos primero y las poblaciones humanas después, respondieron con sucesivas adaptaciones frente a una serie de modificaciones en las circunstancias tanto ambientales como sociales, que condicionaron la evolución de nuestros organismos y de nuestro pensamiento en el sentido en el que efectivamente se produjo.
Por otro lado, es razonable esperar que si las circunstancias ambientales y sociales hubieran sido otras, también podrían haber sido otras las adaptaciones y otras las características de nuestro cerebro y nuestro pensamiento.

Una vez admitida tanto la relatividad como un cierto grado de contingencia de las categorías del conocimiento humano, se pueden distinguir en el momento actual dos posturas mayoritarias, bastante diferentes.
Cuando se las examina se advierte que dependen de presupuestos filosóficos.

Algunas personas niegan la existencia de una realidad objetiva y (aunque sea solo asintóticamente) cognoscible. Adoptan un relativismo evolutivo, moderado, como el trasuntado en la cita de Munévar, a comienzos de este capítulo.
Por lo general, su concepción de la ciencia es la correspondiente al instrumentalismo.

Otras personas continúan adhiriendo al realismo, pero desde una postura perspectivista como la de von Bertalanffy o la de Delbrück (en la cita con la que finaliza el capítulo 1 de este libro). Reconocen que el relativismo presenta límites, y llegan a esa conclusión a partir del examen de los mismos resultados científicos que parecen sustentar las formas más radicales del relativismo.

---

[337] E. Kandell, J. Schwartz y Th. Jessell "Principios de neurociencia", Parte V, Mc Graw-Hill, Madrid, 2001.

16.3 Ciencia, filosofía y concepciones del mundo.

Como vimos en 15.5, las ciencias de la naturaleza y su método de investigación, son la respuesta que las personas de la época moderna pudieron dar enfrentados al objetivo de control sobre la naturaleza, que en esa época pasó a un primer plano.
Las ciencias contemplan la naturaleza y hasta donde les resulta posible, la sociedad, bajo su aspecto de número y medida. Para poder hacerlo, las ciencias se deben auto-limitar en cuatro aspectos fundamentales.

En primer lugar la ciencia establece un corte en el ámbito de los valores, reclamando para sí los valores instrumentales (en el marco de la ciencia se puede estudiar qué es lo que la gente prefiere hacer) y dejando fuera de su alcance los valores normativos (en el marco de la ciencia no corresponde ocuparse de lo que la gente debería preferir hacer).

En segundo lugar establece un corte en el ámbito de los significados, reclamando para sí los significados parciales que aparecen en el contexto de sus teorías específicas y excluyendo los significados globales y los significados existenciales.
Desde una perspectiva científica se puede enfocar bien cualquier parte del dominio de la ciencia. Pero la totalidad de este dominio no se puede captar como un todo desde esa perspectiva.
Asimismo, el punto de vista científico en sí mismo no es suficiente para que un resultado sea considerado relevante por una persona que eventualmente pueda comprenderlo. Siempre puede decidir ignorarlo.

En tercer lugar, la ciencia establece una línea de demarcación entre lo que tiene un propósito y lo que no lo tiene. Enfoca todo lo que estudia asumiendo que no encierra propósito alguno. Como lo estableció claramente Monod [338], la piedra angular del método científico es la negación sistemática de que el verdadero conocimiento se pueda obtener interpretando los fenómenos en términos de algún propósito.

En cuarto lugar, la ciencia elimina de su horizonte la cualidad, porque la cualidad en sí misma no es conmensurable, no se puede reducir a número y medida. Si bien son las cualidades las que se vinculan con los valores normativos, con los significados globales y existenciales, y con los propósitos, las cualidades son vivenciales y subjetivas.
Es un asunto para las disciplinas humanísticas como la literatura, la historia y la filosofía (en su vertiente especulativa) y para las religiones, pero no un asunto para la ciencia.

Con esas inclusiones y esas exclusiones, el pensamiento científico entreteje con sus hilos deductivos la estructura esencial de una trama que, dentro de sus límites, es insuperable: por el grado de acuerdo intersubjetivo que alcanza, por las leyes en las que expresa las regularidades que halla, y por el control que permite alcanzar sobre los aspectos de lo real que ceden a su empuje.

Aún sin detenernos en propósitos, valores, significados y cualidades, basta una consideración cuidadosa de los contenidos propios de las ciencias, para darse cuenta que

[338] J. Monod "*El azar y la necesidad*", Tusquets, Barcelona, 1989.

surgen de efectuar abstracciones, en mayor o menor medida, a partir de los datos de observaciones y experimentos. Pero **abstraer es prescindir**. Por eso el conocimiento científico solamente nos puede dar cuadros incompletos de la realidad, solamente nos puede ofrecer una perspectiva sobre el mundo.
Lo que las ciencias nos ofrecen no es, pues, toda la realidad, sino una perspectiva parcial. En su fuerza (exactitud, reducción a leyes, capacidad de predicción) estriba su debilidad, porque siendo un conocimiento legítimo e imprescindible, siempre será esquemático, parcial y aspectual.

Por otra parte la base empírica de la ciencia no es absoluta: la ciencia no está cimentada sobre roca. Como gustaba decir a Popper, sus teorías se elevan como si fueran una estructura sobre pilotes, en terreno arenoso. Los pilotes se introducen en el terreno. Cuando interrumpimos la hinca de los mismos, no es porque encontramos una roca firme, sino porque nos parece que se encuentran a una profundidad suficiente para soportar la estructura, al menos la estructura que tenemos hasta el momento.
Una idea similar planteó en forma independiente y previa Vaz Ferreira, cuando comparó el conocimiento humano con un mar y a la ciencia como un témpano flotante. El otro tipo de conocimiento de índole discursiva-racional, el conocimiento filosófico, se encuentra en cierto modo en las profundidades alrededor del contorno sumergido del témpano. Complementa al conocimiento científico, lo pone en perspectiva, y dejando paulatinamente de lado las herramientas y los esquemas de las ciencias intenta una aproximación más profunda, más directa, abstracta y flexible, y por ende menos esquemática, menos precisa y como contrapartida ineludible, muchas veces bastante más confusa y discutible.

El conocimiento discursivo-racional es el propio del científico y el del filósofo, por más que al comienzo de muchos procesos intelectuales la intuición juega un rol capital.
Pero en ciencias y en filosofía la intuición debe legitimarse siempre a través de la razón.
Esto sería todo si el ser humano fuera preponderantemente un ser teórico, alejado de la vida, un ser en cuyas venas no corre verdadera sangre sino el humor enrarecido de la razón, considerada como mera actividad intelectual.
Pero es un ser práctico, un ser con emociones y voluntad, un ser de conocimiento y acción.
En la esfera práctica la aprehensión intuitiva de diversos aspectos de la realidad (mundo exterior; yo; valores estéticos, éticos y religiosos) posee una certeza inmediata e irresistible que ninguna inferencia racional parece poder suministrar por sí sola.
Como señala Scheler [339], la religión pertenece a una esfera del valor completamente autónoma: el conocimiento religioso posee autonomía epistemológica, se apoya en la experiencia religiosa, en las vivencias y la conciencia religiosa, y fundamentalmente, en una revelación. Así pues, debido a su carácter parcial, el conocimiento científico y el saber filosófico no excluyen de modo alguno a la fe.
No obstante, muchas personas piensan que es real solo lo que cae dentro del alcance de la ciencia y consideran que el único conocimiento es el conocimiento científico, un conocimiento basado en última instancia en el número y la medida.
"Al ocuparse de un solo plano ontológico, la ciencia puso en cuestión la existencia misma de otros planos y, puesto que ese reto no fue adecuadamente respondido, su visión acabó por imponiendo su impronta al mundo moderno. Por ello la modernidad

[339] M. Scheler "*De lo eterno en el hombre*", Encuentro, Madrid, 2007.

puede definirse como una perspectiva que únicamente contempla la existencia de un solo plano ontológico." [340]

Si bien no parece haber una verdadera concepción del mundo compartida por la mayoría de las personas en nuestra época postmoderna, se ha impuesto progresivamente una imagen física del Universo en expansión a partir de una explosión inicial.
En la base de esta imagen física del mundo se halla el modelo cosmológico estándar (visto en 12.5, 14.5 y 15.5) junto con una articulación de los resultados de la teoría de la relatividad con los de la teoría cuántica.
Esta imagen resulta compatible con muchos puntos de vista diferentes, desde el materialismo emergentista contemporáneo hasta ciertas posiciones religiosas tradicionales.
Los medios audiovisuales han contribuido a consolidarla en narrativas acompañadas de representaciones sugerentes pero a menudo engañosas.

Como dice Ortega y Gasset en la cita al comienzo de este capítulo, la correspondencia entre las observaciones y los conceptos de la física teórica se ha hecho equívoca. Los conceptos físicos ya no son imágenes más o menos directas de las cosas, ya no existe semejanza entre ellos y las cosas: los conceptos físicos forman parte de estructuras simbólicas expresadas en un lenguaje matemático cada vez más abstracto. Apenas se puede exigir que la ordenación de los símbolos pueda expresar la ordenación de los fenómenos.
Lo que se perdió en contenido intuitivo al abandonar la plasticidad de la imagen mecánica del mundo físico propia del siglo XIX [341], se ganó en capacidad de unificar y predecir. Si la palabra ciencia debe continuar aludiendo a conocimiento, entonces, a diferencia de lo que al parecer pensaba Ortega, por conocimiento ya no debe entenderse la presencia en nuestra mente de lo que las cosas son.[342]

---

[340] H. Smith "*La verdad olvidada*", Kairós, Barcelona, 2001, p. 25.

[341] Aún un científico tan eminente como Lord Kelvin, confesaba sentirse incómodo con la teoría electromagnética: "No quedo satisfecho mientras no logro construir un modelo mecánico del objeto estudiado; cuando he conseguido esto, comprendo; de otro modo, no llego a comprender. He aquí por qué no alcanzo a entender la teoría electromagnética de la luz. Aspiro a llegar a comprender, dentro de lo posible, lo que es la luz sin recurrir a cosas que comprendo aún menos. Por ello me atengo a la dinámica pura, porque en ella puedo encontrar modelos, cosa que no sé hacer en el campo de la teoría electromagnética." William Thompson, en "*Lectures on molecular dynamics and the wave theory of light*", J. Clay & Sons, Baltimore, 1904, p. 270.

[342] "Quizás el rasgo más relevante de la ciencia moderna, desde el punto de vista filosófico, es la emergencia de estructuras simbólicas abstractas como el núcleo duro de objetividad detrás-como lo pone Eddington- del cuento colorido de la mente subjetiva, narradora de historias.", H. Weyl y F. Wilczek "*Philosophy of Mathematics and Natural Science*", Princeton University Press, N.J., 2009.

17. Bibliografía

17.1 Temas de matemática

-A. Aleksandrov, A. Kolmogorov, M. Laurentiev y otros "*La matemática: su contenido, métodos y significado*", 3 tomos, Alianza, Madrid, 1973.
-R. Courant y F. John "*Introducción al cálculo y análisis matemático*", 2 tomos, Limusa, México, 2002.
-S.Hasani "*Mathematical Physics*", Springer, New York, 2000.
-G. Korn y Th. Korn "*Matematical handbook for scientists and engineers*", Dover, New York, 2000.
-E.Kreyszig, "Matemáticas Avanzadas para Ingeniería", 2 tomos, Limusa-Wiley, 3ª Edición, México, 2000.
-D.Lovelock y H.Rund "*Tensors, differential forms and variational principles*", Wiley, New York, 1975.
-L. Santaló "*Vectores y tensores con sus aplicaciones*", EUDEBA, Buenos Aires, 1969.

17.2 Temas de física.

-J. Brehm y W. Mullin "*Introduction to the structure of matter: a course in modern physics*", Wiley, N. Y., 1989.
-M. Guerra "*Física: Elementos fundamentales*", 2 tomos, Reverté, Barcelona, 1984.
-G. Fraser (Ed.) "*The new physics in the twenty first century*", Cambridge University Press, U.K, 2009.
-G.Joos e I.Freeman "*Theoretical Physics*", 3ª edición, Dover, NewYork, 1986.
-L. Landau y E. Lifchitz "*Curso abreviado de física teórica*", 2 tomos, Mir, Moscú, 1970.
-L. Landau y E. Lifchitz "*Mecánica*", Reverté, Barcelona, 1981.
-L. Landau y E. Lifchitz "*Teoría del campo*", Reverté, Barcelona, 1981.
-F. Sears, A. Lewis y R. Freedman "*Física universitaria con física moderna*", 2 tomos, Pearson, México, 2005.

17.3 Temas generales relacionados con relatividad y cosmología.

-J. Barrow "*Las constantes de la naturaleza*", Crítica, Barcelona, 2006.
-A. Einstein y otros "*La teoría de la relatividad*", Altaya, Barcelona, 1993.
-B. Greene "*The hidden reality: parallel universes and the deep laws of the cosmos*", Alfred Knopf, N.Y., 2011.
-D. Layzer "*Construcción del Universo*", Labor, Barcelona, 1989.
- R. Penrose *"El camino a la realidad: Una guía completa de las leyes del universo"*, Debate, Madrid, 2006.
-K. Thorne "*Agujeros negros y tiempo curvo*", Crítica, Barcelona, 1995.

-Tanto "*Mundo Científico*" (la traducción de "*La Recherche*") como "*Investigación y Ciencia*" (la traducción de "*Scientific American*") contienen artículos actualizados y escritos a un muy buen nivel de divulgación, sobre el principio de relatividad, las teorías restringida y generalizada, física relativista, física cuántica y de partículas, teorías que intentan unificar la teoría cuántica con la teoría generalizada de la relatividad, astronomía, astrofísica y cosmología.

17.4 Temas de relatividad restringida y generalizada.

- H. Brown, "*Physical Relativity"*, Oxford: Clarendon Press, 2005.
-A.Einstein "*La Relatividad. Memorias originales*", Emecé, Buenos Aires, 1950.
-A. Einstein "*El significado de la relatividad*", Espasa Calpe Argentina, Buenos Aires, 1952.
-A. Einstein, H.Lorentz, H. Minkowski, y H. Weyl "*The principle of relativity*", Dover, New York, 1952.
-E. Gourgoulhon "*Special Relativity in General Frames"*, Berlin, Springer, 2013.
-M. Hobson, G. Efstathiou y A. Lasenby "*General Relativity"*, New York, Cambridge University Press, 2006.
-D. Ivanenko y G. Sardanashvili "*Gravitación*", Editorial URSS, Moscú, 2005.
-E. Loedel "*Física relativista*" Kapelusz, Buenos Aires, 1955.
-W. Mc Crea "*Física relativista*", UTHEA, México, 1965.
-C. Misner, K. Thorne y J. Wheeler "*Gravitation*" W. Freeman, San Francisco, 1973.
-W. Rindler "*Relativity: special, general and cosmological*", Oxford University Press, U.K., 2006.
-H. Schwarz "*Introduction to special relativity*", Mc Graw-Hill, New York, 1968.
-J. L. Synge "*Relativity: the special theory*", North-Holland, Amsterdam, 2ª edición, 1965.
-E.Taylor y J.Wheeler "*Space time physics*", W.Freeman, New York, 1997.
-H. Weyl "*Space, Time, Matter*" Dover, NewYork, 1950.

17.5 Temas de astronomía, astrofísica, relatividad general y cosmología física.

-A. Acker "*Astronomie et astrophysique*", Dunod, París, 2004.
-H. Bondi "*Cosmology*", Dover, New York, 2010.
-J. Comellas "*El mundo de las estrellas*", Equipo Sirius, Madrid, 1999.
-E. Harrison "*Cosmology: the science of the Universe*", Cambridge University Press, Londres, 2008.
-R. Lambourne "*Relativity, gravitation and cosmology*", Cambridge University Press, U.K. 2010.
-A. Liddle "*An introduction to modern cosmology*", Wiley, New York, 2008.
-A. Mazure, G. Mathez y Y. Mellier "*Chronique de l'espace-temps: du vide quantique à l'expansion cosmique*", Masson, Paris, 1997.
-J. Minahan *"Joe´s relatively small book of Special Relativity"*, Uppsala, 2011.
-J. Minahan *"Gravitation och Kosmologi. Lecture Notes"*, Uppsala, 2012.
-D. Raine y E. Thomas "*Black holes: an introduction*", Imperial College Press, Londres, 2005.
-J. Rich "*Fundamentals of Cosmology*", Springer Verlag, Heildelberg, 2001.
-W. Rindler, "*Relativity"*, New York, Oxford University Press, 2006
-S. Serjeant "*Observational cosmology*", Cambridge University Press, U.K, 2010.
-J. Singh "*Teorías de la cosmología moderna*", Alianza, Madrid, 1974.

17.6 Temas de filosofía, historia y metodología de la ciencia.

-J. Barbour "*Absolute or relative motion*?", Volumen 1, "*The discovery of dynamics*", Cambridge University Press, U.K., 1989; Volumen 2, "*The deep structure of general relativity*", Oxford University Press, N.Y., 2007.
-C. Broad "*El pensamiento científico*", Tecnos, Madrid, 1963.

-M. Bunge "*Emergencia y Convergencia*", Gedisa, Barcelona, 2004.
-J. Butterfield y J. Earman "*Philosophy of physics*", Elsevier, N.Y. , 2006.
-E. Cassirer "*Determinism and indeterminism in modern physics: historical and systematic studies of the problem of causality*", Yale University Press, New Haven, USA, 1956.
-E. Cassirer "*El Problema del Conocimiento*", 4 tomos (especialmente el tomo 4, libro primero: *Las ciencias exactas*), Fondo de Cultura Económica, México, 1948-1950.
-W.C.Dampier "*A History of Science and its relation to Philosophy and Religion*", 4ª ed., Cambridge University Press, U.K., 1949.
-M. Delbrück "*Mente y materia: ensayo sobre epistemología evolutiva*", Alianza Editorial, Madrid, 1986.
-H. Kragh "*Higher speculations, grand theories and failed revolutions in physics and cosmology*", Oxford University Press, N.Y., 2011.
-T. Kuhn*"La estructura de las revoluciones científicas"*, FCE, Madrid y México, 2ª edición, 1976.
-J. H. Poincaré "*El valor de la ciencia*", Espasa Calpe, Buenos Aires, 1946.
-J. H. Poincaré "*La ciencia y la hipótesis*", Espasa Calpe, Buenos Aires, 1943.
-K. Popper "*La lógica de la investigación científica*", Tecnos, Madrid, 1962.
-K. Popper "*Conocimiento objetivo. Un enfoque evolucionista*", Tecnos, Madrid, 1998.
-A. Rivadulla "*Éxito, razón y cambio en física: un enfoque instrumental en teoría de la ciencia*", Trotta, Madrid, 2004.
-F. Romero, "*Teoría del hombre*", Losada, Buenos Aires, 1965.
-B.Russell "*La sabiduría de Occidente: Visión histórica de la filosofía occidental en su marco social y político*", Aguilar, Madrid, 1964.
-B. Russell "*El conocimiento humano*", Planeta-Agostini, Barcelona, 1992.
-C. Vaz Ferreira "*Estudios filosóficos. Antología*", Aguilar, Buenos Aires, 1961.
-L.von Bertalanffy "*Teoría General de los Sistemas: Fundamentos, desarrollos y aplicaciones*", Fondo de Cultura Económica, México, 1976.
-C. von Weizsäcker, Th. Görnitz y H. Lyre "*The structure of Physics*", Springer, Dodrecht, Holanda, 2006.
-H. Weyl "*Filosofía de las matemáticas y de la ciencia natural*", Centro de Estudios Filosóficos, UNAM, México,1965.

Con libre acceso en internet se tienen:
-La "Internet Encyclopedia of Philosophy". Ofrece numerosos artículos de interés sobre temas filosóficos variados.
-La "Stanford Encyclopedia of Philosophy". Contiene artículos actualizados sobre el principio de relatividad, sobre las teorías restringida y generalizada y sobre cosmología, entre otros temas tratados en este libro.
-Los artículos de Wikipedia en inglés, sobre una variedad de temas de matemática, física, astronomía, cosmología, epistemología e historia de la ciencia.